PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/62, C07K 13/00, G01N 33/50, C12N 1/19 — A1

(11) International Publication Number: WO 94/25607

(43) International Publication Date: 10 November 1994 (10.11.94)

(21) International Application Number: PCT/US94/04379

(22) International Filing Date: 20 April 1994 (20.04.94)

(30) Priority Data:
08/052,207 23 April 1993 (23.04.93) US

(71) Applicant: UNIVERSITY OF IOWA RESEARCH FOUNDATION [US/US]; Technology Innovation Center, Oakdale, IA 52242 (US).

(72) Inventor: TEEM, John, L.; 1244 E. College, Iowa City, IA 52245 (US).

(74) Agents: SEASE NEBEL, Heidi et al.; Zarley, McKee, Thomte, Voorhees & Sease, Suite 3200, 801 Grand Avenue, Des Moines, IA 50309 (US).

(81) Designated States: AU, CA, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD OF ASSAYING FOR DRUGS WHICH RESTORE FUNCTION OF MUTANT CFTR PROTEIN

(57) Abstract

A method of identifying potentially useful treatments which restore the function of CFTR or other related protein mutations is disclosed. The method involves creation of yeast STE6-CFTR chimeric sequences which upon introduction of a desired mutation inhibit expression of the yeast STE6 gene. A yeast assay (STE6Δ) transformed to include the hybrid chimeras is then exposed to a potential treatment for the mutation. A treatment which restores the function of the STE6 gene and thereby allows mating of the yeast strain is indicative of a potential remedial treatment for the mutation in this selected protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

Title: METHOD OF ASSAYING FOR DRUGS WHICH RESTORE FUNCTION OF MUTANT CFTR PROTEIN

BACKGROUND OF THE INVENTION

Cystic fibrosis is a human genetic disease of the secretory epithelia. Although the survival rate of those suffering with cystic fibrosis has improved in recent years, the median age for patient survival is still only about twenty five to thirty years despite intensive supportive and prophylactic treatment. Today cystic fibrosis remains the most common congenital disease among Caucasians, where it has a prevalence of about one in two thousand live births, and is uniformly fatal. Nearly all patients suffering from the disease develop chronic progressive disease of the respiratory system, the most common cause of death being pulmonary disease. Also, in the majority of cases, pancreatic dysfunction occurs; hepatobiliary and genitourinary diseases are also frequent. Because of the multi-system clinical manifestations of the disease, current methods of treatment for the disease have focused on therapeutic approaches to reduce the symptoms of cystic fibrosis.

United States Patent No. 5,100,647 to Agus, et al, discloses a method for treating cystic fibrosis by administration of the compound sparteine (dodecahydro-7, 14methano-2H, 6H-di-pyrido [1,2-a: 1',2'-e] [1,5] diazocine), acting as a direct exogenous activator of chloride conductants in epithelial airways. United States Patent No. 5,179,001 to Young, et al, discloses a method of treating pulmonary complications associated with cystic fibrosis caused by the gram negative bacterium Pseudomonas aeruginosa. United States

Patent No. 4,826,679 to Roy relates to an oral composition for alleviating digestive manifestation in persons afflicted with cystic fibrosis comprising a therapeutic amount of taurine. Despite much advancement in the treatment of the symptoms of cystic fibrosis very little has been done to effectively "cure" the disease at a genetic level.

One method of gene therapy proposed is United States Patent No. 5,149,797 disclosing a method of site specific alteration of RNA and production of encoded polypeptides. This invention is drawn to correcting the abnormal mRNA present in individual cells, cleaving the mRNA by site directed RNAase followed by introduction of the appropriate oligoribonucleotide followed by endogenous RNA ligase and thus production of a wild-type mRNA encoding a normal protein product which then may be translated to produce the correct protein.

Cystic fibrosis is characterized at the genetic level by mutations in the gene encoding the cystic fibrosis transmembrane conductance regulator (CFTR) protein. The CFTR protein is a member of a superfamily of structurally related membrane proteins named the traffic ATPase or ATP-binding cassette (ABC) transporters. Other members of this superfamily include the multi-drug resistance (MDR) or P-glycoprotein, bovine adenyl cyclase, the yeast STE6 protein, as well as several bacterial amino acid transport proteins (Riordan, et al, 1989; Hyde, et al, 1990). The principle distinguishing feature of this superfamily of proteins is a highly conserved nucleotide-binding domain (NBD).

CFTR is a protein of approximately 1480 amino acids consisting of two repeated motifs, each

comprising 6 transmembrane segments and a nucleotide-binding domain. The two motifs are separated by a large polar R-domain containing multiple potential phosphorylation sites.

Many members of the ABC transport superfamily transport small molecules across biological membranes in an ATP dependent manner. However, CFTR is a Cl^- channel regulated by phosphorylation and by cytosolic ATP (Anderson, et al, "Nucleoside Triphosphates are Required to Open CFTR Chloride Channel" *Cell*,67,775-784). The two nucleotide-binding domains present in CFTR play a central role in the function of CFTR Cl^- channels; they interact directly with ATP to open the CFTR Cl^- channel and ATP hydrolysis may be required for this effect. The importance of the NBD's is further emphasized by the large number of cystic fibrosis associated mutations that have been found in the domains. The most common cystic fibrosis associated mutation, accounting for approximately 68% of the cystic fibrosis chromosomes is deletion of a phenylalanine at position 508 (ΔF508) in the middle of NBD1.

This mutation is characterized by a loss of apical membrane Cl^- channel activity in the epithelial cells leading to chronic progressive disease of the respiratory system. The ΔF508 mutation causes loss of apical membrane Cl^- channel activity affecting the CFTR protein in two ways. First, the mutation causes defective processing and hence mislocalization of the mutant protein. This is presumed to occur from misfolding of the mutant protein so that it fails to exit from the endoplasmic reticulum and progress to the Golgi complex in the apical membrane. The resulting Cl^-

transport defect associated with the ΔF508 is thus largely a consequence of the absence of functional CFTR at the plasma membrane. Second the ΔF508 mutation affects the function of the Cl^- channel. Single channel analysis of CFTR ΔF508 suggests that although the mutant protein is functional, it has an open state probability (P_o) that is one third that of wild-type CFTR.

Defective trafficking of CFTR ΔF508 can be partially reversed if cells are grown at a lower temperature, indicating that mutant protein is still functional. When CFTR ΔF508 is produced at a reduced temperature (less than 37°C) some of the mutant protein transists the Golgi complex normally and is localized correctly in the plasma membrane. (See Denning, G.M.; Anderson, M.P.; Amara, J.; Marshalt, J.; Smith, A.E.; and Welsh, M.J.); "Processing of Mutant CFTR (ΔF508) is Temperature Sensitive." *Nature*,358,761-764. Since CFTR ΔF508 retains partial function, a possible therapy for cystic fibrosis may involve pharmacologic intervention designed to correct the processing defects so that more CFTR ΔF508 reaches the plasma membrane. Potential pharmacologic intervention would represent a non gene-therapy approach for the treatment of cystic fibrosis.

As can be seen, a great need exists in the art for methods of treating cystic fibrosis that do not concentrate solely on the symptoms produced by the disease, but rather treat the disease at a cellular level by attempting to restore function to the mutant product. In light of the scientific potential for pharmacologic treatments, a method for identification of new candidate drugs for cystic

fibrosis treatment is warranted. This is complicated, however, by the fact that a screening method to identify treatments that reverse the CFTR ΔF508 CI^- transport defect in mammalian cells is impractical; large numbers of chemical compounds cannot easily be screened for an activity that restores the cAMP-stimulated Cl^- transport of mutant CFTR.

It is an object of this invention to provide an in vitro method of screening pharmaceutical compositions, and further mutations for their ability to restore function to defective ABC transport proteins resulting from mutations such as CFTR ΔF508, by simple means of a yeast mating assay.

It is another object of the invention to construct chimeric gene sequences which encode the yeast STE6 gene and the CFTR gene for screening of treatments which will restore function to the mutant proteins.

It is another object of the invention to construct chimeric gene sequences which encode the yeast STE6 gene and the P-glycoprotein gene (or other ABC transporter genes) for screening treatments which will inhibit the function of the chimeric protein.

Yet another object of the invention is to provide a method for identifying potentially useful compounds for further useful study in the treatment of disease by means of a yeast mating assay.

Other objects of the invention will become obvious from the description of the invention which follows.

SUMMARY OF THE INVENTION

This invention relates to a method for screening treatments which restore the function of mutant CFTR by simple means of a yeast cell mating assay. According to the present invention, chimeric genetic sequences which encode the yeast a-factor transporter STE6 and CFTR or another member of the related superfamily such as P-glycoprotein are constructed whereby a segment of the CFTR genetic sequence is exchanged for a corresponding segment of the STE6 genetic sequence to produce a product which will complement a yeast ste6Δ mutation but is non-functional upon introduction of a desired mutation into the CFTR sequence.

The resulting chimeras are then transformed into a yeast strain which has the STE6 gene deleted (ste6Δ). The yeast strain is then exposed to the desired drug or treatment to be screened and the strain is assayed for the presence of restoration of mating, which indicates a functional chimera. The invention discloses several chimeric sequences useful for screening for utility of treatments and reversing the effects of the CFTR ΔF508 mutation which is associated with cystic fibrosis.

DETAILED DESCRIPTION OF THE DRAWINGS

Figure 1 represents the amino acid sequence alignment of NBD1 of STE6 and CFTR.

Figure 2 depicts the STE6-CFTR chimeric constructs and their mating efficiency.

Figure 3 is a graphical depiction of a yeast assay according to the present invention which presents a potential cystic fibrosis treatment.

Figure 4a-g is the gene sequence for chimera H4.

Figure 5a-g is the gene sequence for chimera H5.

Figure 6a-g is the gene sequence for chimera H6.

DETAILED DESCRIPTION OF THE INVENTION

The STE6 gene of Saccharomyces cerevisiae encodes an integral membrane protein that functions to transport the peptide a-factor out of the cell during mating. (Kuchler and Thorner, 1992), "Functional expression of human MDR1 in yeast Saccharomyces cerevisiae." *Proc. Natl. Acad. Sch.*, USA89, 2303-2306. Deletion of the STE6 gene from yeast (ste6Δ) results in a sterile phenotype. According to the present invention, chimeric STE6 genes were constructed in which segments of the NBD1 of STE6 were substituted by the analagous NBD1 sequences from CFTR.

Selection for sites of exchange was accomplished by comparing the known amino acid sequences of CFTR and STE6. The CFTR and STE6 amino acid sequences were aligned by use of the "BEST FIT" algorithm for the comparison of DNA and amino acid sequences obtained from the University of Wisconsin and found in the GCG Sequence Analysis Package.

Figure 1 represents an amino acid sequence alignment of the NBD for proteins CFTR and STE6. The CFTR amino acid sequence is depicted on top illustrating amino acids 442 through 578; the corresponding analagous STE6 amino acid sequence is depicted below illustrating amino acids 376 through 536. Corresponding amino acids are indicated by

solid vertical bars and give an indication of potential junctions for the genetic constructs.

Table 1 indicates the portion of CFTR amino acids sequence inserted into STE6 for the different hybrids.

TABLE 1

HYBRID IDENTIFIER	PORTION OF CFTR AMINO ACID SEQUENCE
H1	D443-Y577
H2	D443-F508
H3	F494-F508
H4	F494-I546
H5	F494-L558
H6	F494-Y577

The H1 STE6-CFTR chimeric gene was constructed with a DNA sequence coding for the entire NBD1 of STE6 (from N377 to A535) replaced by the corresponding region of CFTR (D443 to Y557). (See Figure 1.) H2 replaced the amino terminal half of NBD and H6 replaced the carboxy terminal half. The remaining constructs were variations between these two extremes. (See Figure 2.) The construction of plasmids with these sequences used methods which are well known to those of skill in the art however the following description is included merely for purposes of illustration and is not intended to limit in any way the invention.

PLASMID CONSTRUCTIONS

Plasmid RFG416 (a gift from Rick Gaber, Northwestern University) is a single copy CEN plasmid with the selectable marker URA3 and the pUC19 polylinker region. A 6.5 kb SalI-SacI fragment containing the STE6 gene was subcloned from STE6-2μ (a gift from John McGrath, Massachusetts

Institute of Technology) into the SalI and SacI sites of the vector RFG416 to produce plasmid JTS6. The yeast TRP1 gene (on a 0.85 kb BgLl1-EcoR1 DNA fragment) was inserted at nucleotide position 1356 of STE6 NBD1, resulting in JTS6T.

To construct the H1 STE6-CFTR hybrid plasmid, a 441 bp DNA fragment containing the CFTR NBD1 region was synthesized using two STE6-CFTR oligonucleotide primers (primer 1, 5'CCTTCGGAAGCAGTCCTGAAAGATAT-3'; primer 2, 5'-GATGAACAATATCTAGGTATCCAAAA-3') and CFTR cDNA template DNA in a polymerase chain reaction (PCR) (Ho, et al., 1989). PCR reactions were performed with a Temp-Tronic thermal cycler (Barnstead/Thermolyne). Oligonucleotide primers 1 and 2 encoded fusion junctions of STE6 L375 to CFTR K442 and CFTR 1578 to STE6 D537, respectively, found in the H1 STE6-CFTR hybrid gene. The CFTR portion of each primer is underlined. STE6 DNA flanking NBD1 was added to each end of the 441 bp fragment by PCR, resulting in a 867 bp DNA fragment consisting of 419 bp of CFTR DNA (encoding CFTR amino acids K442-L578) that is flanked at the 5' end by 168 bp of STE6 DNA (encoding STE6 amino acids K319-L375) and 280 bp of STE6 DNA at the 3' end (encoding STE6 amino acids D537-G640). Plasmid construction of H1 was performed by cotransformation (Ito, et al, 1983) of yeast strain JPY201 with the 867 bp DNA fragment and 3 μg of plasmid JTS6T and selection of trans-formants on SD-URA (yeast nitrogen base supplemented with all amino acids except uracil). Homologous recombination between the STE6 DNA sequences at each end of the 867 bp DNA fragment with the STE6 gene on the plasmid results in the targeted integration (Orr-Weaver, et al, 1981) of the CFTR sequences into

NBD1 of STE6 on plasmid JTS6T and the consequent loss of the TRP1 gene. Recombinants containing the desired STE6-CFTR hybrid gene were identified as trp auxoptrophs at a frequency of about 1% among the transformants. Plasmid DNA was prepared from trp transformants and the structure of the STE6-CFTR gene was confirmed by DNA sequencing analysis. STE6-CFTR hybrid genes H2-H6 were similarly constructed using the following oligonucleotide primers for constructing the appropriate STE6CFTR junctions; STE6 Q440/CFTR F494, 5'-CACCGTCGTAGAACAGTTTTCCTGGATTA-3'; CFTR G509/STE6 S457, 5'-CCGAATCTGTTGAACCAAAGATGATATTT-3'; CFTR G550/STE6 G509, 5'-TTGTTGTTGCCCGCCACTCAGTGTGATTC-3'; CFTR R560/STE6 A519, 5' ATCTCTGATGAATGCTCTTGCTAAA-GAAAT-3'.

The gene sequences for H4, H5, and H6 are included in Figures 4-6.

The resulting plasmids were then transformed into a yeast strain with the STE6 gene deleted (ste6Δ) to test for complementation by means of a yeast mating assay.

YEAST CELL MATING ASSAY

Several types of variations of assays may be used and are commonly known to those of skill in the art. The following description is not intended to limit in any way the invention.

The yeast strains JPY201 (MATa ste6Δ:HIS3, gal2, ura3-52, lys2-801, trp1, leu2-3,112, his3 Δ200) and 22-2D (MATa, ura352, leu2-3,112,trp1) were used for all mating experiments. JPY201 contains a STE6 deletion (including NBD1 and extending beyond the termination codon) and replacement with the

yeast HIS3 gene (McGrath and Varshavsky, 1989). Quantitative mating assays were performed as in Trueheart, et al (1987).

Transformants of JPY201 containing each STE6-CFTR chimera were grown to log phase in 0.1% glucose SD-URA media. From each culture, 3 x 10^6 cells were mixed with an equal number of 22-2D cells grown in YPD media and collected by filtration onto a Millipore filter, which was then placed upon a YPD plate for 4 hours at 30°C. Cells were resuspended, sonicated briefly, and plated from serial dilutions onto SD+LEU/TRP (yeast nitrogen base supplemented with leucine and tryptophan). Diploid colonies were counted after 3 days at 30°C. Control strains STE6 wild-type and ste6Δ consisted of JPY201 transformed with plasmids JTS6 and JTS6T, respectively. For qualitative petri dish mating assays, JPY201 transformants containing STE6-CFTR chimeras were grown as patches on SD-URA media and then replica printed to YPD media on a lawn of 22-2D cells. Following incubation at 30°C for 8 hr, the plate containing the mating cells was replica printed to SD-LEU/TRP and incubated 3 days at 30°C to allow growth of diploid colonies.

The complementation results are given in Figure 2. As can be seen yeast transformants containing the H1 sequence were unable to complement the ste6Δ mutation, indicating a nonfunctional STE6 chimera. Similarly, the H2 sequence replacing the amino half only was nonfunctional. See Figures 1 and 2. These results suggest that the amino terminal region of NBD1 from CFTR cannot substitute for that of STE6.

In contrast, the STE6-CFTR chimeras containing the central part of NBD1 from CFTR (H3 and H4, See

Figures 1 and 2) maintained a-factor transport activity comparable with the wild-type STE6 yeast strain. This suggests that amino acid residues between the conserved Walker A and B motifs (Walker, et al, 1982) although different in the NBD1 of CFTR and STE6 provide a similar overall structure and function. When the chimeras included progressively larger segments of the carboxy-terminal region of NBD1 from CFTR (H5 and H6) mating efficiency was reduced. Yeast transformed with the hybrid gene H5 which contained the substitution of 74 residues of STE6 NBD1 (from R441 to I516) by that from CFTR (from F494 to L558) mated at 12% efficiency of wild-type STE6. Mating efficiency was further reduced in H6 to background levels. However, a low level of mating was observed when the H6 gene construct was expressed on a high copy number plasmid.

INTRODUCTION OF THE ΔF508 MUTATION

A previous study has suggested that missense mutations in STE6 that were analogous to CF-associated mutations in NBD1 of CFTR result in defective a-factor transport, however, single amino acid deletions analogous to ΔF508 within the central region of STE6 NBD1 had no effect on the STE6 function. (Berkower and Michaelis, "Mutational Analysis of the Yeast a-factor Transporter STE6; a Member of the ATP Binding Cassette (ABC Protein Superfamily)," *EMBO J* 10, 3777-3785 (1991)).

Contrary to this finding, it was found that by incorporating CFTR into NBD1 of the STE6 gene and also including the ΔF508 mutation within the CFTR segment, function of the STE6 chimeric gene was inhibited. The chimeric sequence constructs H3, H4,

H5, and H6 which demonstrated complementation of the ste6Δ yeast strain were also used to assess the effect of the ΔF508 mutation. The ΔF508 mutation was introduced into H3, H4, H5, and H6 (resulting in H3-ΔF508, H4-ΔF508, H5-ΔF508, and H6-ΔF508 respectively). Figure 2 shows the effect of the ΔF508 mutation on a-factor transport as assessed by the quantitative yeast cell mating assay.

As seen from the results in Figure 2, yeast transformed with the H3-ΔF508 chimera mated with an efficiency equal to the H3 control (no F508 mutation). A decrease in mating efficiency (40%) was observed with the strain containing the H4-ΔF508 chimera relative to the H4 control. (Sequence I.D. No. 1). The ΔF508 mutation lowered mating efficiency by 80 fold in the H5-ΔF508 chimera as compared with H5 (Sequence I.D. No. 3). The H6-ΔF508 chimera displayed a comparable decrease in mating as compared to H6 (Sequence I.D. No. 5) efficiency due to the ΔF508 mutation when expressed on a high copy number plasmid, (and relative to the H6 chimera gene also on high copy number plasmid vector). For use in accordance with the present invention, expression of the H6-ΔF508 gene on a high copy number plasmid would be required.

Based on the foregoing junction analysis, the portion of CFTR most promising for substitution into STE6 include the amino acid identity locations between a portion of NBD1 from amino acid 444 up to amino acid 509 through 577. As earlier seen, the segment from 443 to 557 was inactive as was the portion from 443 to 508. Preferred ranges include from amino acid 494 to 577 including: sequence I.D. No. 1, CFTR amino acid F494-I546; Sequence I.D. No.

3, F494-L558; Sequence I.D. No. 5 CFTR amino acid F494-Y577. Several areas of identity between the two aligned sequences as demonstrated in Figure 1 exist within this range of CFTR amino acid substitutions and will represent other chimeric sequences useful for the present invention.

YEAST MATING ASSAY TO IDENTIFY COMPOUNDS THAT CORRECT THE ΔF508 DEFECT

In accordance with the present invention, a yeast strain transformed to contain the discussed chimeric sequences will be used in a yeast mating assay to identify compounds which will correct the affect of the mutation as indicated by an increase in mating. Such an assay may be widely varied by those of skill in the art. Generally, the transformed strain will be combined with a strain of the alpha mating type on a suspension of yeast cells spread on the surface of agar medium in the petri dish. These two strains would be unable to mate due to the ΔF508 mutation.

A compound is then introduced to the medium. For example a plant screening method may be performed by placing a small disc of leaf material removed by means of a paper punch from a plant and placed on the surface of the agar media. Plant compounds will diffuse from the leaf and interact with the yeast strains.

Another method would be to use a filter paper which has been soaked in the desired chemical and then placed on the yeast suspension. Further, other desired mutations can be introduced into the CFTR region to attempt to identify a revertant mutation and a corresponding protein for further study. It is to be understood that any form of exposing the

yeast strain to the desired treatment may be used including modifying external conditions such as temperature.

The mating plate is then incubated for 24 hours to allow yeast cell mating to occur. If a compound corrects the molecular defect associated with H5-ΔF508 transporter then a-factor is transported allowing mating between the two strains, resulting in diploid cells. The mating plate is then replica printed to a selective media plate to allow growth of diploid colonies.

For example strain A could be - MATa, STE6::HIS3, ura3-52,lys2,trpl-298, his3-200, leu2-3,112 <plasmid H5-ΔF508, URA3, CEN>; strain B - MATα, ura3-52, trpl, leu2-3,112.

The selective media used would then be lacking lysine and uracil. Consequently strains A and B could not grow on this media (strain A has the lys2 mutation and strain B has the ura3-52 mutation; these mutations prevent them from going on media without lysine and uracil respectively), however diploids produced as a result of the fusion between strain A and B can grow (diploids contain the genomes of both haploids and consequently has the complementing URA3 and LYS2 wild-type allele).

Production of yeast diploid cells in the area surrounding the treatment disc would be a strong indicator of presence of compound(s) which correct the ΔF508 defect. A number of yeast mating type assays may be developed easily by those of skill in the art and the foregoing description is not intended to limit in any way the invention.

EXAMPLE 1

REVERTANTS OF ΔF508 MATING DEFECT IN YEAST (MUTATION IN R553-L558 REGION TO SUPPRESS ΔF508 IN NBD1)

The R553-L558 amino acid region of CFTR is of great interest in CFTR and other trafficking ATPase's as it is directly adjacent to the conserved LSGGQ linker sequence of NBD1, postulated to function as a transducer of signals between the hydrophobic domains and the NBD's (Mimura, et al, 1991). It was hypothesized that a possible treatment to correct the F508 mutation could be interaction of the region of the NBD polypeptide containing F508, with the R553-L558 region. The only difference between the H5 chimera and the H4 chimera is 6 amino acids in the R553-L558 region. This region is also not present in the H3 chimera in which the ΔF508 mutation had no effect. Possibly there is an interaction between the two regions of the NBD required for correct folding or function (or both) of the NBD. Thus the method of the present invention was used to test whether a mutation in the region R553-L558 might suppress the adverse effects of the ΔF508 in NBD1.

To isolate revertants of the H5-ΔF508 mating defect, the H5-ΔF508 construct was mutagenized in the R553-L558 region in vitro by site-specific oligonucleotide mutagenesis (Ho, et al, "Site-Directed Mutagenesis by Overlap Extension Using the Polymerase Chain Reaction Gene" 77,51-59 (1989). The mutagenesis method was designed to introduce random amino acid substitutions at single codon positions within the R553-L558 region of the H5-ΔF508 plasmid. The procedure allowed 20 possible substitutions at each amino acid position within

R553-L558. The mutagenized DNA was transformed into yeast, and transformant colonies were analyzed by cell mating assays on petri dishes to identify colonies with a mating efficiency higher than the unmutagenized H5-ΔF508 control. Two colonies identified displayed higher efficiencies of mating within the first 20 mutant transformants analyzed. These yeast transformants each contained an H5-ΔF508 plasmid with a mutation at amino acid R553 of CFTR; in one case, R553 was replaced by methionine (H5-ΔF508/R553M), and in the other plasmid, R553 was replaced by glutamine (H5-ΔF508/R553Q). It is possible that other mutations within the R553-L558 region of the H5-F508 plasmid could also result in increased mating efficiency; however, the R553Q and R553M mutants were analyzed in greater detail without further mutagenesis of the R553-L558 region.

The relative mating efficiency of the ΔF508 revertant strains are shown in Table 2, with results expressed relative to the original H5 strain. Whereas the mating efficiency of the H5-ΔF508 yeast strain is approximately 1% of the H5 strain, yeast containing the H5-ΔF508/R553Q and H5-ΔF508/R553M plasmids mated at 3% and 32%, respectively. The revertant mutations, therefore, partially suppress the ΔF508 mating defect. The R553M mutation alone had little effect on H5 (H5-R553M); when this mutant was transformed into yeast, no further increase in mating efficiency was observed as compared with yeast containing H5 (Table 2). In accordance with this invention, mutations were found which suppress the effects of the ΔF508 mutation.

TABLE 2

Relative Mating Efficiency of ΔF508 Reverants

Genotype	Mating Efficiency Relative to H5 (%)
H5-R553M	77.4 = 3.7
H5-ΔF508/R553M	34.2 = 7.8
H5-ΔF508/R553Q	3.2 = 0.8
H5-ΔF508	1.1 = 0.5

Mating efficiencies were determined by quantitative mating assays and are expressed as a percentage relative to H5 (100%). Data are means = SEM.

EXAMPLE 2

TESTING OF THE R553M and R553Q MUTATIONS IN THE MAMMALIAN SYSTEM

To determine whether the mutations which corrected the ΔF508 mutation would also correct the defect in human CFTR ΔF508, the R553Q and R553M mutations were introduced into CFTR ΔF508 cDNA and mammalian cells were transfected with these constructs to determine whether the revertant mutations would correct the defect in cAMP-regulated Cl^- transport of CFTR ΔF508.

The revertants R553Q and R553M were tested for their effects on mislocalization and the channel gating of CFTR ΔF508. The identified revertants modified both of these effects. Mutation of R553 to methionine partially corrected the processing of CFTR ΔF508 as assessed by three criteria; it produced fully glycosylated protein as indicated by electrophoretic separation on SDS polyacrylamide gel; it increased the appearance of the mutant protein in the plasma membrane as determined immunocytochemically; and it produced functional

channels in the plasma membrane when measured by the SPQ halide efflux assay. These results demonstrate the functional utility of the yeast model system for identifying revertants of CF-associated mutations of the NBD1 of CFTR and its use in drug screening for pharmacologic compounds that correct the mating defect. The mutations discovered could be used to identify peptides which may be of utility in reversing the effects of the ΔF508 mutation.

EXAMPLE 3

R553Q AND R553M SUPPRESS THE CFTR ΔF508 ANION TRANSPORT DEFECT

The effect of the R553Q and R553M mutations on CFTR function was determined by a cAMP-stimulated halide efflux using the halide-sensitive fluorophore 6-methoxy-N-(3-sulfopropyl)-quinolinium (SPQ) (Illsley and Verkman, 1987). Expression of CFTR cDNA containing either the R553Q or the R553M mutation alone (without the ΔF508 mutation) in HeLa cells generated cAMP-stimulated halide efflux like wild-type CFTR. Cells expressing the ΔF508 allele in this recombinant system showed little, if any, cAMP-stimulated halide efflux. However, when the mutations R553Q and R553M were introduced into CFTR ΔF508 (CFTR ΔF508/R553Q and CFTR ΔF508/R553M, respectively), cAMP-dependent anion permeability was restored. These results indicate that the CFTR Cl^- channel defect observed with the ΔF508 mutant could be suppressed by either R553 mutation.

Cl^- transport by CFTR ΔF508 containing the R553Q and R553M mutations would be detected only if the processing defect of CFTR ΔF508 was suppressed. To determine further whether the suppressor

mutations correct the processing defect associated with the ΔF508 mutation, the glycosylation patterns of CFTR and the various mutants expressed in HeLa cells were examined. CFTR is a glycoprotein that undergoes progressive glycosylation, resulting in three bands that migrate at different rates on an SDS-polyacrylamide gel. Band A is the most rapidly migrating and represents the nascent, unglycosylated protein; band B has an intermediate rate of migration and a pattern of core glycosylation consistent with processing in the endoplasmic reticulum; band C migrates most slowly and has a pattern of mature glycosylation consistent with processing in the Golgi complex. CFTR ΔF508 was only presented as the unglycosylated band A and the core glycosylated band B protein, consistent with its failure to traverse the Golgi complex and reach the plasma membrane.

We first examined the glycosylation state of CFTR in HeLa cells expressing CFTR cDNA containing either the R553Q and R553M mutations alone (without the ΔF508 mutation). Band C was present in cells expressing wild-type CFTR and also mutant CFTR containing either the R553Q or R553M mutation. In contrast, only bands A and B were present in cells expressing CFTR ΔF508. Thus, the R553Q and R553M mutations alone do not affect the glycosylation of CFTR.

In cells transfected with the CFTR ΔF508/R553M, a small increase in band C was detectable as compared with CFTR ΔF508. This result was observed in three separate experiments. Band C was not able to be consistently detected in cells expressing CFTR ΔF508/R553Q possibly owing to

limitations in the sensitivity of the assay. However, a detectable increase in Band C with CFTR ΔF508/R553M as compared with CFTR ΔF508 occurred.

To assess further the ability of suppressor mutations to correct the mislocalization defect of CFTR ΔF508, immunocytochemistry was used to detect CFTR at the cell surface in HeLa cells expressing wild-type or mutant CFTR. An antibody specific to an external epitope in the first extracellular loop of CFTR was used. (Denning, et al, 1992b). Wild-type CFTR is generally detected at the surface of unpermeabilized HeLa cells. In contrast, no CFTR was detectable at the surface of cells expressing CFTR ΔF508. However, when cells expressed CFTR ΔF508/R553M, CFTR was detected at the plasma membrane. For CFTR ΔF508/R553Q, plasma membrane staining of CFTR was weak and variable and could not be demonstrated consistently. Only nonspecific staining was observed with preimmune serum or in the absence of CFTR. These results were consistent with the observation that more CFTR ΔF508/R553M than CFTR ΔF508/R553Q was found in the band C form and suggest that the amount of CFTR ΔF508/R553Q at the plasma membrane is exceedingly low.

SINGLE-CHANNEL ANALYSIS OF CFTRΔF508/R553Q CFTR

To determine whether the suppressor mutations altered the single-channel properties of CFTR ΔF508, the revertant mutant CFTR ΔF508/R553Q was expressed in HeLa cells and analyzed with the patch-clamp technique. Single-channel analysis of CFTR ΔF508 indicated that the P_o of the CFTR ΔF508 Cl^- channel was reduced as compared with wild-type CFTR. Denning, et al, (1992a) reported that CFTR ΔF508 had

a P_o of 0.13 ± 0.01 (after incubation at reduced temperature), compared with a value of 0.34 ± 0.02 for wild-type, and Dalemans, et al, (1991) reported that the P_o of CFTR ΔF508 Cl^- channels ranged from 0.05 to 0.10, compared with a range of 0.22 to 0.35 for wild type.

As expected from the response observed in the SPQ studies, CFTR/R553Q and CFTR ΔF508/R553Q formed functional Cl^- channels. Inside-out membrane patches from either mutant showed no Cl^- channel activity under basal conditions. However, addition of protein kinase A(PKA) and ATP to the cytosolic surface of the membrane patch activated channels in both cases; both agents were required to activate the channels (n = 6 for each mutant). Single-channel events were recorded from +60 mV to -120 mV in increments of 20 mV. The channels were Cl^- selective as indicated by the reversal potentials (E = 20 mV:E_{Cl^-} = 27 mV). CFTR/R553Q had a single-channel slope conductance (from -120 mV to O mV) of 9.8 ± 0.5 pS (n = 7), and CFTR ΔF508/R553Q had a conductance of 10.2 ± 0.2 pS (n = 11). The current measured at -100 mV for CFTR/R553Q (-1.1 ± 0.1 pA, n = 6) and CFTR ΔF508/R553Q (-1.2 ± 0.0 pA, n = 5) was not different from the current measured for wild-type CFTR (-1.1 ± 0.0 pA, n = 7). Thus, these data indicate that neither mutation altered the single-channel conductive properties of the CFTR Cl^- channel, nor did the mutations abolish the regulation by PKA. The results also indicate that, as predicted by the functional analysis of CFTR ΔF508/R553Q by the SPQ halide efflux assay above, CFTR ΔF508/R553Q is localized in the plasma membrane.

Following activation with PKA and ATP, we measured the P_o for CFTR/R553Q and CFTR ΔF508/R553Q at different MgATP concentrations. As the concentration of MgATP increased, both CFTR/R553Q and CFTR ΔF508/R553Q spent more time in the open state. The P_o of CFTR/R553Q Cl^- channels was similar to that previously reported for wild-type CFTR Cl^- channels (Anderson and Welsh, 1992). However, at 1mM ATP, the P_o of CFTR ΔF508/R553Q (0.29 ± 0.02, $n = 7$) was greater than that of CFTR ΔF508 Cl^- channels (0.13 ± 0.01, $n = 4$). Thus, the R553Q mutation corrected the functional defect in gating of the CFTR ΔF508 Cl^- channels.

YEAST MATING ASSAY WITH PLANT EXTRACT SCREENING

A suspension of yeast cells (strains A and B) was spread upon the surface of agar media in a petri dish.

The suspension consisted of equal proportions of two haploid yeast strains. Strain A - MATa, STE6::HIS3, ura3-52,lys2,trpl-298, his3-200, leu2-3,112 <plasmid H5-ΔF508, URA3, CEN>; strain B -MATα, ura3-52, trpl, leu2-3,112.

Yeast strains A and B are of the a and α mating types respectively, and can mate to form yeast diploids. Strain A contains the STE6/CFTR hybrid a-factor transporter gene H5-ΔF508 (Teem, et al, 1993). The hybrid a-factor transporter encoded by the H5-ΔF508 gene is defective as a result of the ΔF508 mutation. Strain A is therefore unable to transport the mating pheromone a-factor efficiently. Since the transport of the a-factor pheromone is required for yeast cell mating, mating between strains A and B is inefficient.

The agar media is one rich in nutrients (nonselective for the growth of strains A and B) necessary for the optimal mating of yeast.

An extract prepared from the plant genus Grindelia was prepared by grinding fresh Grindelia leaves in a mixture of ethanol and water (1:1). This extract contains the fraction of plant material soluble in a solution of 1:1 ethanol and water. Small filter paper discs were prepared using a standard paper punch and Whatman filter paper #2. A filter paper disc containing the Grindelia extract was prepared by adding 10 µl. of extract to the disc, and then allowing the disc to air dry. Additional extract was added in 10 µl aliquots, allowing the disc to dry after each application until a total of 30 µl. was added to the disc. The disc was then placed upon the surface of the media, on top of the yeast cells.

The mating plate was incubated for 24 hours at 30°C to allow yeast cell mating to occur.

Chemical compounds that diffuse out of the filter paper disc into the agar will come in contact with the mating cells. A component of the Grindelia extract promotes mating of strains A and B.

The mating plate was replica printed to a selective media plate that is incubated at 30°C for three days to allow the growth of diploid colonies.

The selective media is lacking lysine and uracil. Consequently, strains A and B cannot grow on this media (strain A has the lys2 mutation and strain B has the ura3-52 mutation: these mutations prevent them from growing on media without lysine and uracil respectively). However, diploids produced as a result of the fusion between strain A

and B can grow (diploids contain the genomes of both haploids and thus contain the complementing URA3 and LYS2 wild-type alleles). Production of yeast diploid cells in the area surrounding a leaf disc would be a strong indication that the leaf disc contained a compound which corrects the ΔF508 defect. Diploids cells were easily detected by replica printing the yeast cells from the mating plate, onto another petri dish plate containing media which is selective for the growth of the diploid cells. This media will allow the diploid yeast cells to grow and form colonies, while preventing the growth of the haploid (unmated) strain A and B yeast cells. The frequency of diploid colonies produced on the selective plate is an indirect measurement of the ability of strain A to transport a-factor and mate.

A high density of yeast colonies on the selective media plate occurred in the location corresponding to the position where the filter paper disc containing the Grindelia extract occurred on the mating plate. This is clearly distinguishable from the low level of diploid colonies which are produced by the mating of the two yeast strains in the absence of a chemical compound that corrects the ΔF508 defect. (See Figure 3.)

As can be seen in Figure 3 the extract reverses the effect of the ΔF508 mutation indicating its potential for further study.

Thus it can be seen the invention accomplishes all of its objectives.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: University of Iowa Research Foundation.

(ii) TITLE OF INVENTION: Method of Assaying for Drugs which Restore Function of CFTR Mutations

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
(B) STREET: 801 Grand Ave. Suite 3200
(C) CITY: Des Moines
(D) STATE: Iowa
(E) COUNTRY: US
(F) ZIP: 50309

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vii) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US94/
(B) FILING DATE: 20-APR-1994

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/052,207
(B) FILING DATE: 04-APR-1993

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Nebel, Heidi S.
(B) REGISTRATION NUMBER: 37,719
(C) REFERENCE/DOCKET NUMBER: Uirf 93-50

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 515-288-3667
(B) TELEFAX: 515-288-1338

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3840 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Saccharomyces cerevisiae

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..3840

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1317..1488
(D) OTHER INFORMATION: /note= "Substituted analogous sequence from human CFTR gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATG AAC TTT TTA AGT TTT AAG ACT ACA AAA CAC TAT CAC ATT TTC AGG        48
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
  1               5                  10                  15

TAC GTG AAC ATA CGG AAT GAC TAC AGG CTG TTA ATG ATA ATG ATA ATA        96
Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
             20                  25                  30

GGT ACC GTG GCA ACA GGC CTA GTG CCG GCA ATT ACT TCT ATC CTG ACG       144
Gly Thr Val Ala Thr Gly Leu Val Pro Ala Ile Thr Ser Ile Leu Thr
         35                  40                  45

GGC AGA GTG TTC GAT CTA CTA TCA GTT TTC GTG GCT AAT GGG TCA CAT       192
Gly Arg Val Phe Asp Leu Leu Ser Val Phe Val Ala Asn Gly Ser His
     50                  55                  60

CAA GGT TTG TAT TCC CAA CTA GTA CAG AGG TCA ATG GCA GTA ATG GCA       240
Gln Gly Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala
 65                  70                  75                  80

CTT GGT GCG GCT TCT GTG CCA GTA ATG TGG CTT TCT CTA ACA AGT TGG       288
Leu Gly Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp
                 85                  90                  95

ATG CAC ATC GGC GAG AGA CAA GGC TTT AGA ATA CGG TCA CAG ATA TTG       336
Met His Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu
            100                 105                 110
```

```
GAG GCA TAT TTG GAG GAA AAG CCA ATG GAA TGG TAC GAC AAT AAT GAA      384
Glu Ala Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu
        115                 120                 125

AAA TTG TTA GGA GAT TTT ACT CAA ATC AAC AGA TGT GTG GAA GAG CTA      432
Lys Leu Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu
    130                 135                 140

AGA TCA AGC TCC GCA GAG GCA TCA GCC ATA ACT TTC CAG AAT TTA GTT      480
Arg Ser Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val
145                 150                 155                 160

GCA ATA TGT GCG CTT CTG GGG ACG TCA TTC TAC TAT TCT TGG TCA TTA      528
Ala Ile Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu
                165                 170                 175

ACT TTA ATT ATT CTT TGC AGC TCT CCA ATA ATC ACA TTT TTT GCA GTG      576
Thr Leu Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val
            180                 185                 190

GTG TTT TCC AGA ATG ATT CAT GTA TAT TCA GAG AAG GAG AAT TCT GAA      624
Val Phe Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu
        195                 200                 205

ACG AGT AAA GCA GCC CAA TTA CTT ACA TGG TCG ATG AAT GCC GCT CAA      672
Thr Ser Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln
    210                 215                 220

TTA GTG AGA TTA TAT TGT ACA CAA CGT CTA GAA AGG AAA AAA TTC AAG      720
Leu Val Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys
225                 230                 235                 240

GAA ATC ATA CTA AAT TGT AAC ACT TTC TTC ATC AAG AGT TGC TTT TTT      768
Glu Ile Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe
                245                 250                 255

GTT GCT GCA AAC GCT GGG ATC TTG AGA TTT TTG ACG TTG ACT ATG TTT      816
Val Ala Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe
            260                 265                 270

GTA CAG GGA TTC TGG TTT GGT TCC GCA ATG ATC AAA AAG GGC AAG CTG      864
Val Gln Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu
        275                 280                 285

AAC ATT AAC GAT GTA ATC ACT TGC TTC CAT TCA TGC ATT ATG CTG GGC      912
Asn Ile Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly
    290                 295                 300

TCG ACT TTA AAT AAT ACA TTA CAC CAA ATA GTT GTT CTT CAA AAA GGC      960
Ser Thr Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly
305                 310                 315                 320
```

```
GGA GTG GCT ATG GAA AAA ATC ATG ACT CTA TTA AAA GAT GGA TCC AAG
Gly Val Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys
325                 330                 335

CGA AAT CCT TTA AAT AAA ACT GTA GCC CAC CAA TTT CCA CTA GAT TAT      1008
Arg Asn Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr
340                 345                 350

GCC ACC AGT GAT CTA ACA TTT GCT AAT GTT TCG TTT TCT TAT CCA AGC      1056
Ala Thr Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser
355                 360                 365

AGA CCT TCG GAA GCA GTT TTA AAG AAC GTT AGT TTA AAT TTC TCT GCA      1104
Arg Pro Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala
370                 375                 380

GGA CAA TTT ACT TTC ATA GTA GGA AAA TCA GGC TCA GGT AAA TCT ACA      1152
Gly Gln Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr
385                 390                 395                 400

TTA TCC AAC TTA TTA TTA AGG TTC TAC GAT GGC TAT AAT GGA TCG ATA      1200
Leu Ser Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile
405                 410                 415

TCT ATC AAT GGC CAC AAT ATC CAA ACA ATC GAC CAA AAA TTG CTA ATT      1248
Ser Ile Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile
420                 425                 430

GAA AAT ATC ACC GTC GTA GAA CAG TTT TCC TGG ATT ATG CCT GGC ACC      1296
Glu Asn Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr
435                 440                 445

ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC      1344
Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr
450                 455                 460

AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GAC ATC TCC AAG TTT      1392
Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe
465                 470                 475                 480

GCA GAG AAA GAC AAT ATA GTT CTT GGA GAA GGT GGA ATC ACA CTG AGT      1440
Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser
485                 490                 495

GGC GGG CAA CAA CAA AGA GTT GCT ATA GCA CGT GCA TTC ATC AGA GAT      1488
Gly Gly Gln Gln Gln Arg Val Ala Ile Ala Arg Ala Phe Ile Arg Asp
500                 505                 510

ACT CCA ATA TTA TTC TTA GAC GAA GCT GTA TCG GCT CTA GAT ATT GTT      1536
Thr Pro Ile Leu Phe Leu Asp Glu Ala Val Ser Ala Leu Asp Ile Val
515                 520                 525                 1584
```

```
                                                                        1632
CAT CGC AAC CTG TTG ATG AAG GCA ATT AGG CAT TGG AGG AAA GGA AAG
His Arg Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys
    530                 535                 540

                                                                        1680
ACT ACA ATC ATA TTG ACG CAT GAG TTG AGC CAA ATT GAA TCT GAT GAC
Thr Thr Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp
545                 550                 555                 560

                                                                        1728
TAT TTA TAT TTA ATG AAG GAA GGT GAA GTT GTT GAA AGC GGC ACC CAG
Tyr Leu Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln
                565                 570                 575

                                                                        1776
TCT GAA CTT CTA GCC GAT CCG ACC ACT ACA TTT AGC ACA TGG TAT CAC
Ser Glu Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His
            580                 585                 590

                                                                        1824
CTA CAG AAT GAC TAC TCT GAT GCG AAA ACT ATT GTA GAT ACA GAG ACT
Leu Gln Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr
        595                 600                 605

                                                                        1872
GAA GAA AAA TCT ATA CAC ACT GTG GAA AGT TTT AAC TCT CAA TTG GAA
Glu Glu Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu
    610                 615                 620

                                                                        1920
ACA CCA AAA CTT GGA TCA TGC TTA AGT AAT CTG GGA TAT GAT GAG ACA
Thr Pro Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr
625                 630                 635                 640

                                                                        1968
GAT CAG TTG TCC TTT TAC GAA GCA ATC TAT CAA AAA AGA TCG AAC GTT
Asp Gln Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val
                645                 650                 655

                                                                        2016
AGA ACA AGA AGG GTT AAA GTT GAA GAG GAA AAT ATT GGG TAT GCA CTA
Arg Thr Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu
            660                 665                 670

                                                                        2064
AAA CAA CAA AAG AAC ACC GAA AGT TCA ACA GGG CCA CAA CTT CTG AGC
Lys Gln Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser
        675                 680                 685

                                                                        2112
ATT ATT CAG ATT ATC AAA AGA ATG ATT AAA AGC ATA AGA TAC AAA AAA
Ile Ile Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys
    690                 695                 700

                                                                        2160
ATT CTA ATC TTG GGA CTG CTA TGT TCT CTT ATC GCA GGC GCC ACA AAT
Ile Leu Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn
705                 710                 715                 720

                                                                        2208
CCC GTC TTT TCA TAC ACA TTC AGT TTC TTA CTA GAA GGA ATT GTC CCA
Pro Val Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro
                725                 730                 735
```

```
TCC ACG GAT GGA AAA ACT GGC TCT TCA CAT TAT TTG GCG AAA TGG TCG      2256
Ser Thr Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser
            740                 745                 750

CTT CTT GTT CTT GGT GTG GCT GCG GCA GAT GGT ATT TTC AAT TTT GCT      2304
Leu Leu Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala
        755                 760                 765

AAA GGA TTC CTA TTA GAT TGC TGC AGT GAA TAC TGG GTT ATG GAT CTT      2352
Lys Gly Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu
    770                 775                 780

AGA AAT GAA GTT ATG GAA AAA CTG ACG AGA AAG AAT ATG GAC TGG TTT      2400
Arg Asn Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe
785                 790                 795                 800

TCT GGT GAA AAC AAC AAG GCT TCT GAA ATT TCT GCT CTA GTC TTG AAT      2448
Ser Gly Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn
                805                 810                 815

GAT TTG CGA GAT TTG AGG TCT TTG GTC TCT GAA TTT TTG AGT GCA ATG      2496
Asp Leu Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met
            820                 825                 830

ACT AGT TTC GTT ACC GTA TCA ACG ATT GGA CTA ATT TGG GCG TTA GTA      2544
Thr Ser Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val
        835                 840                 845

TCG GGC TGG AAG TTA AGT TTG GTT TGT ATT TCG ATG TTT CCA CTC ATA      2592
Ser Gly Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile
    850                 855                 860

ATT ATA TTT TCA GCA ATA TAT GGA GGC ATT TTA CAA AAA TGC GAA ACA      2640
Ile Ile Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr
865                 870                 875                 880

GAT TAT AAG ACA TCT GTT GCT CAG TTA GAA AAC TGC CTG TAC CAG ATT      2688
Asp Tyr Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile
                885                 890                 895

GTC ACT AAC ATT AAA ACC ATT AAG TGC TTA CAA GCT GAA TTT CAT TTT      2736
Val Thr Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe
            900                 905                 910

CAA TTG ACC TAC CAT GAC TTG AAG ATA AAA ATG CAA CAA ATT GCC TCC      2784
Gln Leu Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser
        915                 920                 925

AAA CGC GCC ATT GCC ACA GGA TTT GGT ATA TCT ATG ACA AAC ATG ATT      2832
Lys Arg Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile
    930                 935                 940
```

```
                                                                        2880
GTC ATG TGT ATC CAA GCT ATT ATT TAC TAC TAT GGC CTA AAG CTG GTT
Val Met Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val
945                 950                 955                 960

                                                                        2928
ATG ATT CAC GAG TAC ACC TCA AAG GAA ATG TTT ACG ACT TTC ACT TTG
Met Ile His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu
                965                 970                 975

                                                                        2976
TTA TTA TTC ACT ATT ATG TCA TGC ACT AGC CTA GTA AGT CAG ATA CCC
Leu Leu Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro
            980                 985                 990

                                                                        3024
GAT ATA AGT AGA GGC CAA CGT GCT GCC AGT TGG ATC TAT AGG ATT CTT
Asp Ile Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu
        995                 1000                1005

                                                                        3072
GAT GAA AAG CAT AAT ACC CTA GAG GTT GAA AAC AAT AAT GCT AGA ACA
Asp Glu Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr
    1010                1015                1020

                                                                        3120
GTG GGA ATA GCT GGT CAC ACC TAC CAT GGC AAA GAA AAA AAA CCA ATC
Val Gly Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile
1025                1030                1035                1040

                                                                        3168
GTT TCA ATT CAA AAT TTG ACA TTT GCC TAT CCA TCT GCA CCT ACC GCC
Val Ser Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala
                1045                1050                1055

                                                                        3216
TTT GTT TAC AAA AAC ATG AAT TTT GAC ATG TTT TGC GGA CAG ACG TTA
Phe Val Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu
            1060                1065                1070

                                                                        3264
GGT ATC ATT GGT GAA TCA GGC ACA GGA AAG TCT ACA CTT GTG CTT TTA
Gly Ile Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu
        1075                1080                1085

                                                                        3312
TTA ACA AAA CTT TAT AAT TGT GAA GTA GGC AAA ATT AAA ATA GAC GGT
Leu Thr Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly
    1090                1095                1100

                                                                        3360
ACG GAC GTA AAT GAC TGG AAT TTG ACA AGT TTA AGA AAA GAA ATT TCA
Thr Asp Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser
1105                1110                1115                1120

                                                                        3408
GTG GTT GAG CAA AAA CCT TTA TTA TTC AAT GGA ACC ATC AGA GAT AAC
Val Val Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn
                1125                1130                1135

                                                                        3456
CTA ACT TAT GGT TTA CAA GAT GAA ATA CTT GAA ATT GAA ATG TAT GAT
Leu Thr Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp
            1140                1145                1150
```

```
GCA TTA AAA TAC GTA GGA ATC CAT GAC TTT GTA ATT TCA TCA CCT CAG      3504
Ala Leu Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln
      1155                1160                1165

GGC TTG GAT ACA CGT ATT GAT ACA ACT TTA CTG TCA GGT GGA CAA GCG      3552
Gly Leu Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala
    1170                1175                1180

CAA AGG CTT TGC ATA GCC AGA GCA CTT CTG AGA AAA TCA AAA ATT CTG      3600
Gln Arg Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu
1185                1190                1195                1200

ATT TTA GAT GAG TGT ACT TCA GCC TTG GAT TCT GTC AGC TCC TCT ATC      3648
Ile Leu Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile
                1205                1210                1215

ATC AAT GAG ATC GTC AAA AAA GGT CCA CCT GCT CTA CTA ACA ATG GTT      3696
Ile Asn Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val
            1220                1225                1230

ATA ACG CAT AGT GAA CAA ATG ATG AGG TCT TGT AAC TCG ATT GCA GTT      3744
Ile Thr His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val
        1235                1240                1245

CTT AAA GAT GGT AAA GTG GTT GAG CGA GGT AAC TTC GAC ACT TTA TAT      3792
Leu Lys Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr
    1250                1255                1260

AAT AAT CGC GGG GAA TTA TTC CAA ATT GTT TCC AAC CAA AGC AGT TAA      3840
Asn Asn Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser
1265                1270                1275
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1279 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
 1               5                  10                  15

Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
            20                  25                  30

Gly Thr Val Ala Thr Gly Leu Val Pro Ala Ile Thr Ser Ile Leu Thr
        35                  40                  45
```

Gly Arg Val Phe Asp Leu Leu Ser Val Phe Val Ala Asn Gly Ser His
50 55 60

Gln Gly Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala
65 70 75 80

Leu Gly Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp
85 90 95

Met His Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu
100 105 110

Glu Ala Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu
115 120 125

Lys Leu Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu
130 135 140

Arg Ser Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val
145 150 155 160

Ala Ile Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu
165 170 175

Thr Leu Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val
180 185 190

Val Phe Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu
195 200 205

Thr Ser Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln
210 215 220

Leu Val Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys
225 230 235 240

Glu Ile Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe
245 250 255

Val Ala Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe
260 265 270

Val Gln Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu
275 280 285

Asn Ile Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly
290 295 300

Ser Thr Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly
305 310 315 320

Gly Val Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys
325 330 335

Arg Asn Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr
340 345 350

Ala Thr Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser
355 360 365

Arg Pro Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala
370 375 380

Gly Gln Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr
385 390 395 400

Leu Ser Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile
405 410 415

Ser Ile Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile
420 425 430

Glu Asn Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr
435 440 445

Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr
450 455 460

Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe
465 470 475 480

Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser
485 490 495

Gly Gly Gln Gln Gln Arg Val Ala Ile Ala Arg Ala Phe Ile Arg Asp
500 505 510

Thr Pro Ile Leu Phe Leu Asp Glu Ala Val Ser Ala Leu Asp Ile Val
515 520 525

His Arg Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys
530 535 540

Thr Thr Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp
545 550 555 560

Tyr Leu Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln
565 570 575

Ser Glu Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His
580 585 590

```
Leu Gln Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr
        595                 600                 605

Glu Glu Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu
    610                 615                 620

Thr Pro Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr
625                 630                 635                 640

Asp Gln Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val
                645                 650                 655

Arg Thr Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu
            660                 665                 670

Lys Gln Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser
        675                 680                 685

Ile Ile Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys
    690                 695                 700

Ile Leu Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn
705                 710                 715                 720

Pro Val Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro
                725                 730                 735

Ser Thr Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser
            740                 745                 750

Leu Leu Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala
        755                 760                 765

Lys Gly Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu
    770                 775                 780

Arg Asn Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe
785                 790                 795                 800

Ser Gly Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn
                805                 810                 815

Asp Leu Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met
            820                 825                 830

Thr Ser Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val
        835                 840                 845

Ser Gly Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile
    850                 855                 860
```

Ile Ile Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr
865 870 875 880

Asp Tyr Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile
885 890 895

Val Thr Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe
900 905 910

Gln Leu Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser
915 920 925

Lys Arg Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile
930 935 940

Val Met Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val
945 950 955 960

Met Ile His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu
965 970 975

Leu Leu Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro
980 985 990

Asp Ile Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu
995 1000 1005

Asp Glu Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr
1010 1015 1020

Val Gly Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile
1025 1030 1035 1040

Val Ser Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala
1045 1050 1055

Phe Val Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu
1060 1065 1070

Gly Ile Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu
1075 1080 1085

Leu Thr Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly
1090 1095 1100

Thr Asp Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser
1105 1110 1115 1120

Val Val Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn
1125 1130 1135

Leu Thr Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp
1140 1145 1150

Ala Leu Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln
1155 1160 1165

Gly Leu Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala
1170 1175 1180

Gln Arg Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu
1185 1190 1195 1200

Ile Leu Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile
1205 1210 1215

Ile Asn Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val
1220 1225 1230

Ile Thr His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val
1235 1240 1245

Leu Lys Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr
1250 1255 1260

Asn Asn Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser
1265 1270 1275

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3840 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Saccharomyces cerevisiae

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..3840

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1318..1521

(D) OTHER INFORMATION: /note= "Substituted analogous sequence from human CFTR gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATG AAC TTT TTA AGT TTT AAG ACT ACA AAA CAC TAT CAC ATT TTC AGG          48
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
 1               5                  10                  15

TAC GTG AAC ATA CGG AAT GAC TAC AGG CTG TTA ATG ATA ATG ATA ATA          96
Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
            20                  25                  30

GGT ACC GTG GCA ACA GGC CTA GTG CCG GCA ATT ACT TCT ATC CTG ACG         144
Gly Thr Val Ala Thr Gly Leu Val Pro Ala Ile Thr Ser Ile Leu Thr
        35                  40                  45

GGC AGA GTG TTC GAT CTA CTA TCA GTT TTC GTG GCT AAT GGG TCA CAT         192
Gly Arg Val Phe Asp Leu Leu Ser Val Phe Val Ala Asn Gly Ser His
    50                  55                  60

CAA GGT TTG TAT TCC CAA CTA GTA CAG AGG TCA ATG GCA GTA ATG GCA         240
Gln Gly Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala
65                  70                  75                  80

CTT GGT GCG GCT TCT GTG CCA GTA ATG TGG CTT TCT CTA ACA AGT TGG         288
Leu Gly Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp
                85                  90                  95

ATG CAC ATC GGC GAG AGA CAA GGC TTT AGA ATA CGG TCA CAG ATA TTG         336
Met His Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu
            100                 105                 110

GAG GCA TAT TTG GAG GAA AAG CCA ATG GAA TGG TAC GAC AAT AAT GAA         384
Glu Ala Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu
        115                 120                 125

AAA TTG TTA GGA GAT TTT ACT CAA ATC AAC AGA TGT GTG GAA GAG CTA         432
Lys Leu Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu
    130                 135                 140

AGA TCA AGC TCC GCA GAG GCA TCA GCC ATA ACT TTC CAG AAT TTA GTT         480
Arg Ser Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val
145                 150                 155                 160

GCA ATA TGT GCG CTT CTG GGG ACG TCA TTC TAC TAT TCT TGG TCA TTA         528
Ala Ile Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu
                165                 170                 175
```

```
                                                                        576
ACT TTA ATT ATT CTT TGC AGC TCT CCA ATA ATC ACA TTT TTT GCA GTG
Thr Leu Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val
        180                 185                 190

                                                                        624
GTG TTT TCC AGA ATG ATT CAT GTA TAT TCA GAG AAG GAG AAT TCT GAA
Val Phe Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu
    195                 200                 205

                                                                        672
ACG AGT AAA GCA GCC CAA TTA CTT ACA TGG TCG ATG AAT GCC GCT CAA
Thr Ser Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln
210                 215                 220

                                                                        720
TTA GTG AGA TTA TAT TGT ACA CAA CGT CTA GAA AGG AAA AAA TTC AAG
Leu Val Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys
225                 230                 235                 240

                                                                        768
GAA ATC ATA CTA AAT TGT AAC ACT TTC TTC ATC AAG AGT TGC TTT TTT
Glu Ile Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe
                245                 250                 255

                                                                        816
GTT GCT GCA AAC GCT GGG ATC TTG AGA TTT TTG ACG TTG ACT ATG TTT
Val Ala Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe
            260                 265                 270

                                                                        864
GTA CAG GGA TTC TGG TTT GGT TCC GCA ATG ATC AAA AAG GGC AAG CTG
Val Gln Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu
        275                 280                 285

                                                                        912
AAC ATT AAC GAT GTA ATC ACT TGC TTC CAT TCA TGC ATT ATG CTG GGC
Asn Ile Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly
    290                 295                 300

                                                                        960
TCG ACT TTA AAT AAT ACA TTA CAC CAA ATA GTT GTT CTT CAA AAA GGC
Ser Thr Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly
305                 310                 315                 320

                                                                       1008
GGA GTG GCT ATG GAA AAA ATC ATG ACT CTA TTA AAA GAT GGA TCC AAG
Gly Val Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys
                325                 330                 335

                                                                       1056
CGA AAT CCT TTA AAT AAA ACT GTA GCC CAC CAA TTT CCA CTA GAT TAT
Arg Asn Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr
            340                 345                 350

                                                                       1104
GCC ACC AGT GAT CTA ACA TTT GCT AAT GTT TCG TTT TCT TAT CCA AGC
Ala Thr Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser
        355                 360                 365

                                                                       1152
AGA CCT TCG GAA GCA GTT TTA AAG AAC GTT AGT TTA AAT TTC TCT GCA
Arg Pro Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala
    370                 375                 380
```

```
GGA CAA TTT ACT TTC ATA GTA GGA AAA TCA GGC TCA GGT AAA TCT ACA                1200
Gly Gln Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr
385                 390                 395                 400

TTA TCC AAC TTA TTA TTA AGG TTC TAC GAT GGC TAT AAT GGA TCG ATA                1248
Leu Ser Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile
                405                 410                 415

TCT ATC AAT GGC CAC AAT ATC CAA ACA ATC GAC CAA AAA TTG CTA ATT                1296
Ser Ile Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile
            420                 425                 430

GAA AAT ATC ACC GTC GTA GAA CAG TTT TCC TGG ATT ATG CCT GGC ACC                1344
Glu Asn Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr
        435                 440                 445

ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC                1392
Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr
    450                 455                 460

AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GAC ATC TCC AAG TTT                1440
Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe
465                 470                 475                 480

GCA GAG AAA GAC AAT ATA GTT CTT GGA GAA GGT GGA ATC ACA CTG AGT                1488
Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser
                485                 490                 495

GGA GGT CAA CGA GCA AGA ATT TCT TTA GCA AGA GCA TTC ATC AGA GAT                1536
Gly Gly Gln Arg Ala Arg Ile Ser Leu Ala Arg Ala Phe Ile Arg Asp
            500                 505                 510

ACT CCA ATA TTA TTC TTA GAC GAA GCT GTA TCG GCT CTA GAT ATT GTT                1584
Thr Pro Ile Leu Phe Leu Asp Glu Ala Val Ser Ala Leu Asp Ile Val
        515                 520                 525

CAT CGC AAC CTG TTG ATG AAG GCA ATT AGG CAT TGG AGG AAA GGA AAG                1632
His Arg Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys
    530                 535                 540

ACT ACA ATC ATA TTG ACG CAT GAG TTG AGC CAA ATT GAA TCT GAT GAC                1680
Thr Thr Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp
545                 550                 555                 560

TAT TTA TAT TTA ATG AAG GAA GGT GAA GTT GTT GAA AGC GGC ACC CAG                1728
Tyr Leu Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln
                565                 570                 575

TCT GAA CTT CTA GCC GAT CCG ACC ACT ACA TTT AGC ACA TGG TAT CAC                1776
Ser Glu Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His
            580                 585                 590
```

```
CTA CAG AAT GAC TAC TCT GAT GCG AAA ACT ATT GTA GAT ACA GAG ACT      1824
Leu Gln Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr
        595                 600                 605

GAA GAA AAA TCT ATA CAC ACT GTG GAA AGT TTT AAC TCT CAA TTG GAA      1872
Glu Glu Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu
    610                 615                 620

ACA CCA AAA CTT GGA TCA TGC TTA AGT AAT CTG GGA TAT GAT GAG ACA      1920
Thr Pro Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr
625                 630                 635                 640

GAT CAG TTG TCC TTT TAC GAA GCA ATC TAT CAA AAA AGA TCG AAC GTT      1968
Asp Gln Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val
                645                 650                 655

AGA ACA AGA AGG GTT AAA GTT GAA GAG GAA AAT ATT GGG TAT GCA CTA      2016
Arg Thr Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu
            660                 665                 670

AAA CAA CAA AAG AAC ACC GAA AGT TCA ACA GGG CCA CAA CTT CTG AGC      2064
Lys Gln Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser
        675                 680                 685

ATT ATT CAG ATT ATC AAA AGA ATG ATT AAA AGC ATA AGA TAC AAA AAA      2112
Ile Ile Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys
    690                 695                 700

ATT CTA ATC TTG GGA CTG CTA TGT TCT CTT ATC GCA GGC GCC ACA AAT      2160
Ile Leu Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn
705                 710                 715                 720

CCC GTC TTT TCA TAC ACA TTC AGT TTC TTA CTA GAA GGA ATT GTC CCA      2208
Pro Val Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro
                725                 730                 735

TCC ACG GAT GGA AAA ACT GGC TCT TCA CAT TAT TTG GCG AAA TGG TCG      2256
Ser Thr Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser
            740                 745                 750

CTT CTT GTT CTT GGT GTG GCT GCG GCA GAT GGT ATT TTC AAT TTT GCT      2304
Leu Leu Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala
        755                 760                 765

AAA GGA TTC CTA TTA GAT TGC TGC AGT GAA TAC TGG GTT ATG GAT CTT      2352
Lys Gly Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu
    770                 775                 780

AGA AAT GAA GTT ATG GAA AAA CTG ACG AGA AAG AAT ATG GAC TGG TTT      2400
Arg Asn Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe
785                 790                 795                 800
```

```
TCT GGT GAA AAC AAC AAG GCT TCT GAA ATT TCT GCT CTA GTC TTG AAT      2448
Ser Gly Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn
                805                 810                 815

GAT TTG CGA GAT TTG AGG TCT TTG GTC TCT GAA TTT TTG AGT GCA ATG      2496
Asp Leu Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met
            820                 825                 830

ACT AGT TTC GTT ACC GTA TCA ACG ATT GGA CTA ATT TGG GCG TTA GTA      2544
Thr Ser Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val
        835                 840                 845

TCG GGC TGG AAG TTA AGT TTG GTT TGT ATT TCG ATG TTT CCA CTC ATA      2592
Ser Gly Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile
    850                 855                 860

ATT ATA TTT TCA GCA ATA TAT GGA GGC ATT TTA CAA AAA TGC GAA ACA      2640
Ile Ile Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr
865                 870                 875                 880

GAT TAT AAG ACA TCT GTT GCT CAG TTA GAA AAC TGC CTG TAC CAG ATT      2688
Asp Tyr Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile
                885                 890                 895

GTC ACT AAC ATT AAA ACC ATT AAG TGC TTA CAA GCT GAA TTT CAT TTT      2736
Val Thr Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe
            900                 905                 910

CAA TTG ACC TAC CAT GAC TTG AAG ATA AAA ATG CAA CAA ATT GCC TCC      2784
Gln Leu Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser
        915                 920                 925

AAA CGC GCC ATT GCC ACA GGA TTT GGT ATA TCT ATG ACA AAC ATG ATT      2832
Lys Arg Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile
    930                 935                 940

GTC ATG TGT ATC CAA GCT ATT ATT TAC TAC TAT GGC CTA AAG CTG GTT      2880
Val Met Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val
945                 950                 955                 960

ATG ATT CAC GAG TAC ACC TCA AAG GAA ATG TTT ACG ACT TTC ACT TTG      2928
Met Ile His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu
                965                 970                 975

TTA TTA TTC ACT ATT ATG TCA TGC ACT AGC CTA GTA AGT CAG ATA CCC      2976
Leu Leu Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro
            980                 985                 990

GAT ATA AGT AGA GGC CAA CGT GCT GCC AGT TGG ATC TAT AGG ATT CTT      3024
Asp Ile Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu
        995                 1000                1005
```

```
GAT GAA AAG CAT AAT ACC CTA GAG GTT GAA AAC AAT AAT GCT AGA ACA      3072
Asp Glu Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr
    1010                1015                1020

GTG GGA ATA GCT GGT CAC ACC TAC CAT GGC AAA GAA AAA AAA CCA ATC      3120
Val Gly Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile
1025                1030                1035                1040

GTT TCA ATT CAA AAT TTG ACA TTT GCC TAT CCA TCT GCA CCT ACC GCC      3168
Val Ser Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala
                1045                1050                1055

TTT GTT TAC AAA AAC ATG AAT TTT GAC ATG TTT TGC GGA CAG ACG TTA      3216
Phe Val Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu
            1060                1065                1070

GGT ATC ATT GGT GAA TCA GGC ACA GGA AAG TCT ACA CTT GTG CTT TTA      3264
Gly Ile Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu
        1075                1080                1085

TTA ACA AAA CTT TAT AAT TGT GAA GTA GGC AAA ATT AAA ATA GAC GGT      3312
Leu Thr Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly
    1090                1095                1100

ACG GAC GTA AAT GAC TGG AAT TTG ACA AGT TTA AGA AAA GAA ATT TCA      3360
Thr Asp Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser
1105                1110                1115                1120

GTG GTT GAG CAA AAA CCT TTA TTA TTC AAT GGA ACC ATC AGA GAT AAC      3408
Val Val Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn
                1125                1130                1135

CTA ACT TAT GGT TTA CAA GAT GAA ATA CTT GAA ATT GAA ATG TAT GAT      3456
Leu Thr Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp
            1140                1145                1150

GCA TTA AAA TAC GTA GGA ATC CAT GAC TTT GTA ATT TCA TCA CCT CAG      3504
Ala Leu Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln
        1155                1160                1165

GGC TTG GAT ACA CGT ATT GAT ACA ACT TTA CTG TCA GGT GGA CAA GCG      3552
Gly Leu Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala
    1170                1175                1180

CAA AGG CTT TGC ATA GCC AGA GCA CTT CTG AGA AAA TCA AAA ATT CTG      3600
Gln Arg Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu
1185                1190                1195                1200

ATT TTA GAT GAG TGT ACT TCA GCC TTG GAT TCT GTC AGC TCC TCT ATC      3648
Ile Leu Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile
                1205                1210                1215
```

ATC AAT GAG ATC GTC AAA AAA GGT CCA CCT GCT CTA CTA ACA ATG GTT 3696
Ile Asn Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val
1220 1225 1230

ATA ACG CAT AGT GAA CAA ATG ATG AGG TCT TGT AAC TCG ATT GCA GTT 3744
Ile Thr His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val
1235 1240 1245

CTT AAA GAT GGT AAA GTG GTT GAG CGA GGT AAC TTC GAC ACT TTA TAT 3792
Leu Lys Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr
1250 1255 1260

AAT AAT CGC GGG GAA TTA TTC CAA ATT GTT TCC AAC CAA AGC AGT TAA 3840
Asn Asn Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser
1265 1270 1275

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1279 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
1 5 10 15

Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
20 25 30

Gly Thr Val Ala Thr Gly Leu Val Pro Ala Ile Thr Ser Ile Leu Thr
35 40 45

Gly Arg Val Phe Asp Leu Leu Ser Val Phe Val Ala Asn Gly Ser His
50 55 60

Gln Gly Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala
65 70 75 80

Leu Gly Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp
85 90 95

Met His Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu
100 105 110

Glu Ala Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu
115 120 125

Lys Leu Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu
130 135 140

Arg Ser Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val
145 150 155 160

Ala Ile Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu
165 170 175

Thr Leu Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val
180 185 190

Val Phe Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu
195 200 205

Thr Ser Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln
210 215 220

Leu Val Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys
225 230 235 240

Glu Ile Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe
245 250 255

Val Ala Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe
260 265 270

Val Gln Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu
275 280 285

Asn Ile Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly
290 295 300

Ser Thr Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly
305 310 315 320

Gly Val Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys
325 330 335

Arg Asn Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr
340 345 350

Ala Thr Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser
355 360 365

Arg Pro Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala
370 375 380

Gly Gln Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr
385 390 395 400

Leu Ser Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile
405 410 415

Ser Ile Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile
420 425 430

Glu Asn Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr
435 440 445

Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr
450 455 460

Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe
465 470 475 480

Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser
485 490 495

Gly Gly Gln Arg Ala Arg Ile Ser Leu Ala Arg Ala Phe Ile Arg Asp
500 505 510

Thr Pro Ile Leu Phe Leu Asp Glu Ala Val Ser Ala Leu Asp Ile Val
515 520 525

His Arg Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys
530 535 540

Thr Thr Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp
545 550 555 560

Tyr Leu Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln
565 570 575

Ser Glu Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His
580 585 590

Leu Gln Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr
595 600 605

Glu Glu Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu
610 615 620

Thr Pro Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr
625 630 635 640

Asp Gln Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val
645 650 655

Arg Thr Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu
660 665 670

Lys Gln Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser
675 680 685

Ile Ile Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys
690 695 700

Ile Leu Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn
705 710 715 720

Pro Val Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro
725 730 735

Ser Thr Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser
740 745 750

Leu Leu Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala
755 760 765

Lys Gly Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu
770 775 780

Arg Asn Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe
785 790 795 800

Ser Gly Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn
805 810 815

Asp Leu Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met
820 825 830

Thr Ser Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val
835 840 845

Ser Gly Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile
850 855 860

Ile Ile Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr
865 870 875 880

Asp Tyr Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile
885 890 895

Val Thr Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe
900 905 910

Gln Leu Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser
915 920 925

Lys Arg Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile
930 935 940

Val Met Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val
945 950 955 960

Met Ile His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu
965 970 975

Leu Leu Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro
980 985 990

Asp Ile Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu
995 1000 1005

Asp Glu Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr
1010 1015 1020

Val Gly Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile
1025 1030 1035 1040

Val Ser Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala
1045 1050 1055

Phe Val Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu
1060 1065 1070

Gly Ile Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu
1075 1080 1085

Leu Thr Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly
1090 1095 1100

Thr Asp Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser
1105 1110 1115 1120

Val Val Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn
1125 1130 1135

Leu Thr Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp
1140 1145 1150

Ala Leu Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln
1155 1160 1165

Gly Leu Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala
1170 1175 1180

Gln Arg Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu
1185 1190 1195 1200

Ile Leu Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile
1205 1210 1215

```
Ile Asn Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val
            1220                1225                1230

Ile Thr His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val
        1235                1240                1245

Leu Lys Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr
    1250                1255                1260

Asn Asn Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser
1265                1270                1275
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3786 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Saccharomyces cerevisiae

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..3786

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1264..1524
(D) OTHER INFORMATION: /note= "Substituted analogous sequence from human CFTR gene"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
ATG AAC TTT TTA AGT TTT AAG ACT ACA AAA CAC TAT CAC ATT TTC AGG      48
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
  1               5                  10                  15

TAC GTG AAC ATA CGG AAT GAC TAC AGG CTG TTA ATG ATA ATG ATA ATA      96
Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
            20                  25                  30
```

```
GGT ACC GTG GCA CTA TCA GTT TTC GTG GCT AAT GGG TCA CAT CAA GGT        144
Gly Thr Val Ala Leu Ser Val Phe Val Ala Asn Gly Ser His Gln Gly
         35                  40                  45

TTG TAT TCC CAA CTA GTA CAG AGG TCA ATG GCA GTA ATG GCA CTT GGT        192
Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala Leu Gly
     50                  55                  60

GCG GCT TCT GTG CCA GTA ATG TGG CTT TCT CTA ACA AGT TGG ATG CAC        240
Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp Met His
 65                  70                  75                  80

ATC GGC GAG AGA CAA GGC TTT AGA ATA CGG TCA CAG ATA TTG GAG GCA        288
Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu Glu Ala
                 85                  90                  95

TAT TTG GAG GAA AAG CCA ATG GAA TGG TAC GAC AAT AAT GAA AAA TTG        336
Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu Lys Leu
            100                 105                 110

TTA GGA GAT TTT ACT CAA ATC AAC AGA TGT GTG GAA GAG CTA AGA TCA        384
Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu Arg Ser
        115                 120                 125

AGC TCC GCA GAG GCA TCA GCC ATA ACT TTC CAG AAT TTA GTT GCA ATA        432
Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val Ala Ile
    130                 135                 140

TGT GCG CTT CTG GGG ACG TCA TTC TAC TAT TCT TGG TCA TTA ACT TTA        480
Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu Thr Leu
145                 150                 155                 160

ATT ATT CTT TGC AGC TCT CCA ATA ATC ACA TTT TTT GCA GTG GTG TTT        528
Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val Val Phe
                165                 170                 175

TCC AGA ATG ATT CAT GTA TAT TCA GAG AAG GAG AAT TCT GAA ACG AGT        576
Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu Thr Ser
            180                 185                 190

AAA GCA GCC CAA TTA CTT ACA TGG TCG ATG AAT GCC GCT CAA TTA GTG        624
Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln Leu Val
        195                 200                 205

AGA TTA TAT TGT ACA CAA CGT CTA GAA AGG AAA AAA TTC AAG GAA ATC        672
Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys Glu Ile
    210                 215                 220

ATA CTA AAT TGT AAC ACT TTC TTC ATC AAG AGT TGC TTT TTT GTT GCT        720
Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe Val Ala
225                 230                 235                 240
```

```
GCA AAC GCT GGG ATC TTG AGA TTT TTG ACG TTG ACT ATG TTT GTA CAG          768
Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe Val Gln
245                 250                 255

GGA TTC TGG TTT GGT TCC GCA ATG ATC AAA AAG GGC AAG CTG AAC ATT          816
Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu Asn Ile
            260                 265                 270

AAC GAT GTA ATC ACT TGC TTC CAT TCA TGC ATT ATG CTG GGC TCG ACT          864
Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly Ser Thr
        275                 280                 285

TTA AAT AAT ACA TTA CAC CAA ATA GTT GTT CTT CAA AAA GGC GGA GTG          912
Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly Gly Val
    290                 295                 300

GCT ATG GAA AAA ATC ATG ACT CTA TTA AAA GAT GGA TCC AAG CGA AAT          960
Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys Arg Asn
305                 310                 315                 320

CCT TTA AAT AAA ACT GTA GCC CAC CAA TTT CCA CTA GAT TAT GCC ACC         1008
Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr Ala Thr
                325                 330                 335

AGT GAT CTA ACA TTT GCT AAT GTT TCG TTT TCT TAT CCA AGC AGA CCT         1056
Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser Arg Pro
            340                 345                 350

TCG GAA GCA GTT TTA AAG AAC GTT AGT TTA AAT TTC TCT GCA GGA CAA         1104
Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala Gly Gln
        355                 360                 365

TTT ACT TTC ATA GTA GGA AAA TCA GGC TCA GGT AAA TCT ACA TTA TCC         1152
Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr Leu Ser
    370                 375                 380

AAC TTA TTA TTA AGG TTC TAC GAT GGC TAT AAT GGA TCG ATA TCT ATC         1200
Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile Ser Ile
385                 390                 395                 400

AAT GGC CAC AAT ATC CAA ACA ATC GAC CAA AAA TTG CTA ATT GAA AAT         1248
Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile Glu Asn
                405                 410                 415

ATC ACC GTC GTA GAA CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA         1296
Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr Ile Lys
            420                 425                 430

GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC AGA AGC         1344
Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr Arg Ser
        435                 440                 445
```

```
GTC ATC AAA GCA TGC CAA CTA GAA GAG GAC ATC TCC AAG TTT GCA GAG                1392
Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe Ala Glu
    450                 455                 460

AAA GAC AAT ATA GTT CTT GGA GAA GGT GGA ATC ACA CTG AGT GGA GGT               1440
Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser Gly Gly
465                 470                 475                 480

CAA CGA GCA AGA ATT TCT TTA GCA AGA GCA GTA TAC AAA GAT GCT GAT               1488
Gln Arg Ala Arg Ile Ser Leu Ala Arg Ala Val Tyr Lys Asp Ala Asp
                485                 490                 495

TTG TAT TTA TTA GAC TCT CCT TTT GGA TAC CTA GAT ATT GTT CAT CGC               1536
Leu Tyr Leu Leu Asp Ser Pro Phe Gly Tyr Leu Asp Ile Val His Arg
            500                 505                 510

AAC CTG TTG ATG AAG GCA ATT AGG CAT TGG AGG AAA GGA AAG ACT ACA               1584
Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys Thr Thr
        515                 520                 525

ATC ATA TTG ACG CAT GAG TTG AGC CAA ATT GAA TCT GAT GAC TAT TTA               1632
Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp Tyr Leu
    530                 535                 540

TAT TTA ATG AAG GAA GGT GAA GTT GTT GAA AGC GGC ACC CAG TCT GAA               1680
Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln Ser Glu
545                 550                 555                 560

CTT CTA GCC GAT CCG ACC ACT ACA TTT AGC ACA TGG TAT CAC CTA CAG               1728
Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His Leu Gln
                565                 570                 575

AAT GAC TAC TCT GAT GCG AAA ACT ATT GTA GAT ACA GAG ACT GAA GAA               1776
Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr Glu Glu
            580                 585                 590

AAA TCT ATA CAC ACT GTG GAA AGT TTT AAC TCT CAA TTG GAA ACA CCA               1824
Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu Thr Pro
        595                 600                 605

AAA CTT GGA TCA TGC TTA AGT AAT CTG GGA TAT GAT GAG ACA GAT CAG               1872
Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr Asp Gln
    610                 615                 620

TTG TCC TTT TAC GAA GCA ATC TAT CAA AAA AGA TCG AAC GTT AGA ACA               1920
Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val Arg Thr
625                 630                 635                 640

AGA AGG GTT AAA GTT GAA GAG GAA AAT ATT GGG TAT GCA CTA AAA CAA               1968
Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu Lys Gln
                645                 650                 655
```

```
CAA AAG AAC ACC GAA AGT TCA ACA GGG CCA CAA CTT CTG AGC ATT ATT      2016
Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser Ile Ile
            660                 665                 670

CAG ATT ATC AAA AGA ATG ATT AAA AGC ATA AGA TAC AAA AAA ATT CTA      2064
Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys Ile Leu
        675                 680                 685

ATC TTG GGA CTG CTA TGT TCT CTT ATC GCA GGC GCC ACA AAT CCC GTC      2112
Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn Pro Val
    690                 695                 700

TTT TCA TAC ACA TTC AGT TTC TTA CTA GAA GGA ATT GTC CCA TCC ACG      2160
Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro Ser Thr
705                 710                 715                 720

GAT GGA AAA ACT GGC TCT TCA CAT TAT TTG GCG AAA TGG TCG CTT CTT      2208
Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser Leu Leu
                725                 730                 735

GTT CTT GGT GTG GCT GCG GCA GAT GGT ATT TTC AAT TTT GCT AAA GGA      2256
Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala Lys Gly
            740                 745                 750

TTC CTA TTA GAT TGC TGC AGT GAA TAC TGG GTT ATG GAT CTT AGA AAT      2304
Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu Arg Asn
        755                 760                 765

GAA GTT ATG GAA AAA CTG ACG AGA AAG AAT ATG GAC TGG TTT TCT GGT      2352
Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe Ser Gly
    770                 775                 780

GAA AAC AAC AAG GCT TCT GAA ATT TCT GCT CTA GTC TTG AAT GAT TTG      2400
Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn Asp Leu
785                 790                 795                 800

CGA GAT TTG AGG TCT TTG GTC TCT GAA TTT TTG AGT GCA ATG ACT AGT      2448
Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met Thr Ser
                805                 810                 815

TTC GTT ACC GTA TCA ACG ATT GGA CTA ATT TGG GCG TTA GTA TCG GGC      2496
Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val Ser Gly
            820                 825                 830

TGG AAG TTA AGT TTG GTT TGT ATT TCG ATG TTT CCA CTC ATA ATT ATA      2544
Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile Ile Ile
        835                 840                 845

TTT TCA GCA ATA TAT GGA GGC ATT TTA CAA AAA TGC GAA ACA GAT TAT      2592
Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr Asp Tyr
    850                 855                 860
```

```
AAG ACA TCT GTT GCT CAG TTA GAA AAC TGC CTG TAC CAG ATT GTC ACT                2640
Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile Val Thr
865                 870                 875                 880

AAC ATT AAA ACC ATT AAG TGC TTA CAA GCT GAA TTT CAT TTT CAA TTG                2688
Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe Gln Leu
                885                 890                 895

ACC TAC CAT GAC TTG AAG ATA AAA ATG CAA CAA ATT GCC TCC AAA CGC                2736
Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser Lys Arg
        900                 905                 910

GCC ATT GCC ACA GGA TTT GGT ATA TCT ATG ACA AAC ATG ATT GTC ATG                2784
Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile Val Met
    915                 920                 925

TGT ATC CAA GCT ATT ATT TAC TAC TAT GGC CTA AAG CTG GTT ATG ATT                2832
Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val Met Ile
930                 935                 940

CAC GAG TAC ACC TCA AAG GAA ATG TTT ACG ACT TTC ACT TTG TTA TTA                2880
His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu Leu Leu
945                 950                 955                 960

TTC ACT ATT ATG TCA TGC ACT AGC CTA GTA AGT CAG ATA CCC GAT ATA                2928
Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro Asp Ile
                965                 970                 975

AGT AGA GGC CAA CGT GCT GCC AGT TGG ATC TAT AGG ATT CTT GAT GAA                2976
Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu Asp Glu
        980                 985                 990

AAG CAT AAT ACC CTA GAG GTT GAA AAC AAT AAT GCT AGA ACA GTG GGA                3024
Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr Val Gly
    995                 1000                1005

ATA GCT GGT CAC ACC TAC CAT GGC AAA GAA AAA AAA CCA ATC GTT TCA                3072
Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile Val Ser
1010                1015                1020

ATT CAA AAT TTG ACA TTT GCC TAT CCA TCT GCA CCT ACC GCC TTT GTT                3120
Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala Phe Val
1025                1030                1035                1040

TAC AAA AAC ATG AAT TTT GAC ATG TTT TGC GGA CAG ACG TTA GGT ATC                3168
Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu Gly Ile
                1045                1050                1055

ATT GGT GAA TCA GGC ACA GGA AAG TCT ACA CTT GTG CTT TTA TTA ACA                3216
Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu Leu Thr
        1060                1065                1070
```

```
AAA CTT TAT AAT TGT GAA GTA GGC AAA ATT AAA ATA GAC GGT ACG GAC      3264
Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly Thr Asp
        1075                1080                1085

GTA AAT GAC TGG AAT TTG ACA AGT TTA AGA AAA GAA ATT TCA GTG GTT      3312
Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser Val Val
    1090                1095                1100

GAG CAA AAA CCT TTA TTA TTC AAT GGA ACC ATC AGA GAT AAC CTA ACT      3360
Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn Leu Thr
1105                1110                1115                1120

TAT GGT TTA CAA GAT GAA ATA CTT GAA ATT GAA ATG TAT GAT GCA TTA      3408
Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp Ala Leu
                1125                1130                1135

AAA TAC GTA GGA ATC CAT GAC TTT GTA ATT TCA TCA CCT CAG GGC TTG      3456
Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln Gly Leu
            1140                1145                1150

GAT ACA CGT ATT GAT ACA ACT TTA CTG TCA GGT GGA CAA GCG CAA AGG      3504
Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala Gln Arg
        1155                1160                1165

CTT TGC ATA GCC AGA GCA CTT CTG AGA AAA TCA AAA ATT CTG ATT TTA      3552
Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu Ile Leu
    1170                1175                1180

GAT GAG TGT ACT TCA GCC TTG GAT TCT GTC AGC TCC TCT ATC ATC AAT      3600
Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile Ile Asn
1185                1190                1195                1200

GAG ATC GTC AAA AAA GGT CCA CCT GCT CTA CTA ACA ATG GTT ATA ACG      3648
Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val Ile Thr
                1205                1210                1215

CAT AGT GAA CAA ATG ATG AGG TCT TGT AAC TCG ATT GCA GTT CTT AAA      3696
His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val Leu Lys
            1220                1225                1230

GAT GGT AAA GTG GTT GAG CGA GGT AAC TTC GAC ACT TTA TAT AAT AAT      3744
Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr Asn Asn
        1235                1240                1245

CGC GGG GAA TTA TTC CAA ATT GTT TCC AAC CAA AGC AGT TAA              3786
Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser
    1250                1255                1260
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1261 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
1 5 10 15

Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
20 25 30

Gly Thr Val Ala Leu Ser Val Phe Val Ala Asn Gly Ser His Gln Gly
35 40 45

Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala Leu Gly
50 55 60

Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp Met His
65 70 75 80

Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu Glu Ala
85 90 95

Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu Lys Leu
100 105 110

Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu Arg Ser
115 120 125

Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val Ala Ile
130 135 140

Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu Thr Leu
145 150 155 160

Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val Val Phe
165 170 175

Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu Thr Ser
180 185 190

Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln Leu Val
195 200 205

Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys Glu Ile
210 215 220

Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe Val Ala
225 230 235 240

Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe Val Gln
245 250 255

Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu Asn Ile
260 265 270

Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly Ser Thr
275 280 285

Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly Gly Val
290 295 300

Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys Arg Asn
305 310 315 320

Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr Ala Thr
325 330 335

Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser Arg Pro
340 345 350

Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala Gly Gln
355 360 365

Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr Leu Ser
370 375 380

Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile Ser Ile
385 390 395 400

Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile Glu Asn
405 410 415

Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr Ile Lys
420 425 430

Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr Arg Ser
435 440 445

Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe Ala Glu
450 455 460

Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser Gly Gly
465 470 475 480

Gln Arg Ala Arg Ile Ser Leu Ala Arg Ala Val Tyr Lys Asp Ala Asp
485 490 495

Leu Tyr Leu Leu Asp Ser Pro Phe Gly Tyr Leu Asp Ile Val His Arg
500 505 510

Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys Thr Thr
515 520 525

Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp Tyr Leu
530 535 540

Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln Ser Glu
545 550 555 560

Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His Leu Gln
565 570 575

Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr Glu Glu
580 585 590

Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu Thr Pro
595 600 605

Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr Asp Gln
610 615 620

Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val Arg Thr
625 630 635 640

Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu Lys Gln
645 650 655

Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser Ile Ile
660 665 670

Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys Ile Leu
675 680 685

Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn Pro Val
690 695 700

Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro Ser Thr
705 710 715 720

Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser Leu Leu
725 730 735

Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala Lys Gly
740 745 750

Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu Arg Asn
755 760 765

Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe Ser Gly
770 775 780

Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn Asp Leu
785 790 795 800

Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met Thr Ser
805 810 815

Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val Ser Gly
820 825 830

Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile Ile Ile
835 840 845

Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr Asp Tyr
850 855 860

Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile Val Thr
865 870 875 880

Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe Gln Leu
885 890 895

Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser Lys Arg
900 905 910

Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile Val Met
915 920 925

Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val Met Ile
930 935 940

His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu Leu Leu
945 950 955 960

Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro Asp Ile
965 970 975

Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu Asp Glu
980 985 990

Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr Val Gly
995 1000 1005

Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile Val Ser
1010 1015 1020

Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala Phe Val
1025 1030 1035 1040

Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu Gly Ile
1045 1050 1055

```
Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu Leu Thr
            1060                1065                1070

Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly Thr Asp
        1075                1080                1085

Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser Val Val
    1090                1095                1100

Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn Leu Thr
1105                1110                1115                1120

Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp Ala Leu
                1125                1130                1135

Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln Gly Leu
            1140                1145                1150

Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala Gln Arg
        1155                1160                1165

Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu Ile Leu
    1170                1175                1180

Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile Ile Asn
1185                1190                1195                1200

Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val Ile Thr
                1205                1210                1215

His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val Leu Lys
            1220                1225                1230

Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr Asn Asn
        1235                1240                1245

Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser
    1250                1255                1260
```

What is claimed is:

1. A method of identifying potential treatments which correct mislocalization and/or misfunction of CFTR resulting from mutations said method comprising: forming a chimeric gene sequence by substituting a portion of a gene sequence which encodes the first nucleotide binding domain of cystic fibrosis conductance regulator with a gene sequence which encodes STE6; said CFTR sequence being a CFTR gene sequence encoding a plurality of amino acids from amino acid 444 to and including amino acid 577, such that introduction of a mutation in the CFTR region inhibits expression of the chimeric STE6 gene producti transforming said chimeric gene sequence into a yeast strain with no functional STE6 gene; introducing said potential treatment to said yeast strain; and assaying for the production of diploid yeast colonies following yeast cell mating.

2. The method of claim 1 wherein said mutation is a deletion of phenylalanine at position 508 of the CFTR gene product.

3. The method of claim 1 wherein said CFTR substitute is the gene sequence which encodes from amino acid 494 through 546 of the CFTR gene product (SEQ ID #1).

4. The method of claim 1 wherein said CFTR substitution is the gene sequence which encodes from amino acid 494 up to and including amino acid 558 of the CFTR gene product (SEQ ID #3).

5. The method of claim 1 wherein said CFTR substitution is the gene sequence which encodes from amino acid 494 up to and including 577 of the CFTR gene product (SEQ ID #5).

6. The method of claim 3, 4, or 5 wherein said CFTR substitution includes a mutation which causes cystic fibrosis.

7. A chimeric protein comprising a yeast STE6 protein product having substituted therein an analogous portion of CFTR nucleotide binding domain including a plurality of amino acids from amino acid 443 to amino acid 577.

8. The protein of claim 7 wherein said CFTR sequence includes a mutation which results in cystic fibrosis.

9. The protein of claim 7 wherein said amino acid sequence is SEQ ID # 2.

10. The protein of claim 7 wherein said amino acid sequence is SEQ ID # 4.

11. The protein of claim 7 wherein said amino acid sequence is SEQ ID # 6.

12. A chimeric protein sufficiently homologous to that of claim 7 such that introduction of a mutation in the CFTR portion will prevent transport of a-factor out a yeast cell thereby preventing yeast mating.

13. A DNA sequence which encodes the protein of claim 7.

14. A host cell transformed with the DNA sequence of claim 13 in a manner allowing expression of the STE6 gene.

15. A method of screening for potential inhibitors of ABC transporters by the use of STE6 chimeras, said method comprising selecting from the class of related proteins consisting essentially of multi-drug resistance (mdr) or P-glycoprotein, the choroquine resistance transporter (pfmdr) of Plasmodium falciparum, and CFTR; substituting the nucleotide sequence of the STE6 gene with the sequence encoding the analogous portion of said ABC transporter; transforming said chimeric gene sequence into a yeast which is deleted for the STE6 gene; introducing a treatment to said transformed yeast strain; and assaying for the production of yeast diploid colonies following yeast cell mating.

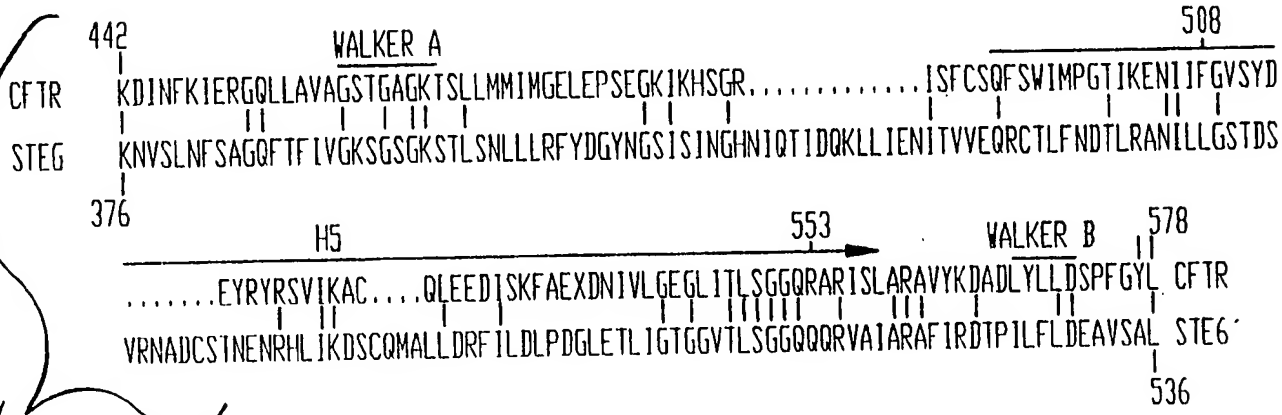
442
WALKER A
508
CFTR KDINFKIERGQLLAVAGSTGAGKTSLLMMIMGELEPSEGKIKHSGR.....ISFCSQFSWIMPGTIKENIIFGVSYD
STEG KNVSLNFSAGQFTFIVGKSGSGKSTLSNLLLRFYDGYNGSISINGHNIQTIDQKLLIENITVVEQRCTLFNDTLRANILLGSTDS
376
H5
553
WALKER B
578
.....EYRYRSVIKAC....QLEEDISKFAEKDNIVLGEGLITLSGGQRARISLARAVYKDADLYLLDSPFGYL CFTR
VRNADCSTNENRHLIKDSCQMALLDRFILDLPDGLETLIGTGGVTLSGGQQQRVAIARAFIRDTPILFLDEAVSAL STE6
536

FIG. 1

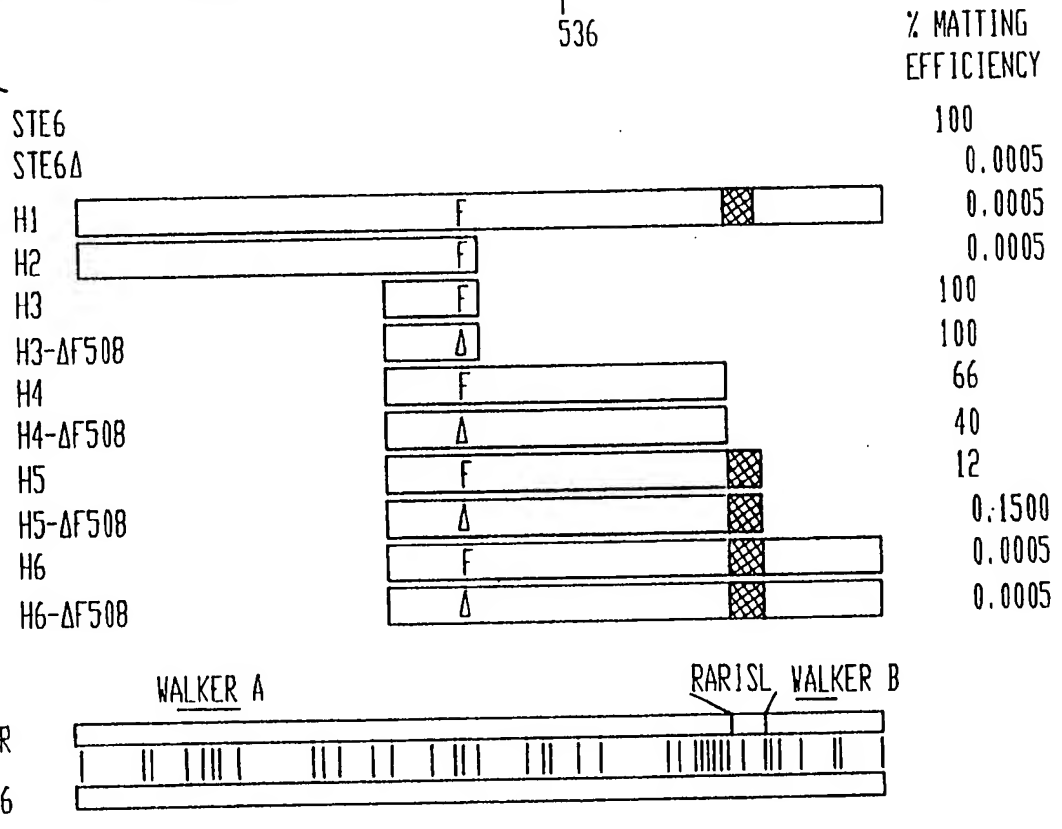
% MATTING EFFICIENCY
STE6 100
STE6Δ 0.0005
H1 F 0.0005
H2 F 0.0005
H3 F 100
H3-ΔF508 Δ 100
H4 F 66
H4-ΔF508 Δ 40
H5 F 12
H5-ΔF508 Δ 0.1500
H6 F 0.0005
H6-ΔF508 Δ 0.0005
WALKER A
RARISL
WALKER B
CFTR
STE6

FIG. 2

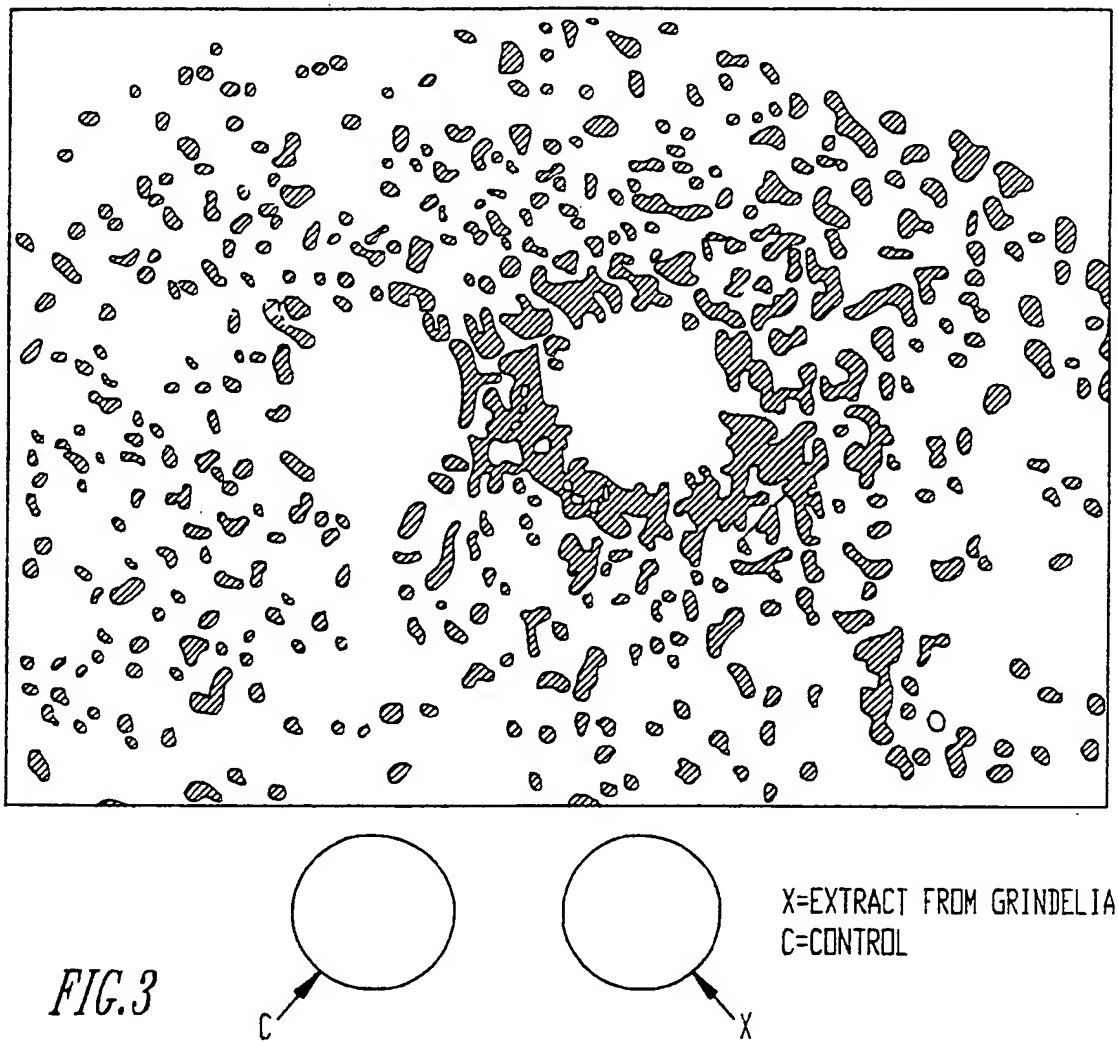

FIG.3

SELECTIVE MEDIA PLATE WITH DIPLOID COLONIES

H4 Chimera (Seq. ID No. 1)

```
ATG AAC TTT TTA AGT TTT AAG ACT ACA AAA CAC TAT CAC ATT TTC AGG        48
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
 1               5                  10                  15

TAC GTG AAC ATA CGG AAT GAC TAC AGG CTG TTA ATG ATA ATG ATA ATA        96
Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
            20                  25                  30

GGT ACC GTG GCA ACA GGC CTA GTG CCG GCA ATT ACT TCT ATC CTG ACG       144
Gly Thr Val Ala Thr Gly Leu Val Pro Ala Ile Thr Ser Ile Leu Thr
        35                  40                  45

GGC AGA GTG TTC GAT CTA CTA TCA GTT TTC GTG GCT AAT GGG TCA CAT       192
Gly Arg Val Phe Asp Leu Leu Ser Val Phe Val Ala Asn Gly Ser His
    50                  55                  60

CAA GGT TTG TAT TCC CAA CTA GTA CAG AGG TCA ATG GCA GTA ATG GCA       240
Gln Gly Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala
65                  70                  75                  80

CTT GGT GCG GCT TCT GTG CCA GTA ATG TGG CTT TCT CTA ACA AGT TGG       288
Leu Gly Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp
                85                  90                  95

ATG CAC ATC GGC GAG AGA CAA GGC TTT AGA ATA CGG TCA CAG ATA TTG       336
Met His Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu
            100                 105                 110

GAG GCA TAT TTG GAG GAA AAG CCA ATG GAA TGG TAC GAC AAT AAT GAA       384
Glu Ala Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu
        115                 120                 125

AAA TTG TTA GGA GAT TTT ACT CAA ATC AAC AGA TGT GTG GAA GAG CTA       432
Lys Leu Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu
    130                 135                 140

AGA TCA AGC TCC GCA GAG GCA TCA GCC ATA ACT TTC CAG AAT TTA GTT       480
Arg Ser Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val
145                 150                 155                 160

GCA ATA TGT GCG CTT CTG GGG ACG TCA TTC TAC TAT TCT TGG TCA TTA       528
Ala Ile Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu
                165                 170                 175

ACT TTA ATT ATT CTT TGC AGC TCT CCA ATA ATC ACA TTT TTT GCA GTG       576
Thr Leu Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val
            180                 185                 190
```

FIG. 4A

```
GTG TTT TCC AGA ATG ATT CAT GTA TAT TCA GAG AAG GAG AAT TCT GAA        624
Val Phe Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu
        195                 200                 205

ACG AGT AAA GCA GCC CAA TTA CTT ACA TGG TCG ATG AAT GCC GCT CAA        672
Thr Ser Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln
    210                 215                 220

TTA GTG AGA TTA TAT TGT ACA CAA CGT CTA GAA AGG AAA AAA TTC AAG        720
Leu Val Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys
225                 230                 235                 240

GAA ATC ATA CTA AAT TGT AAC ACT TTC TTC ATC AAG AGT TGC TTT TTT        768
Glu Ile Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe
                245                 250                 255

GTT GCT GCA AAC GCT GGG ATC TTG AGA TTT TTG ACG TTG ACT ATG TTT        816
Val Ala Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe
            260                 265                 270

GTA CAG GGA TTC TGG TTT GGT TCC GCA ATG ATC AAA AAG GGC AAG CTG        864
Val Gln Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu
        275                 280                 285

AAC ATT AAC GAT GTA ATC ACT TGC TTC CAT TCA TGC ATT ATG CTG GGC        912
Asn Ile Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly
    290                 295                 300

TCG ACT TTA AAT AAT ACA TTA CAC CAA ATA GTT GTT CTT CAA AAA GGC        960
Ser Thr Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly
305                 310                 315                 320

GGA GTG GCT ATG GAA AAA ATC ATG ACT CTA TTA AAA GAT GGA TCC AAG       1008
Gly Val Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys
                325                 330                 335

CGA AAT CCT TTA AAT AAA ACT GTA GCC CAC CAA TTT CCA CTA GAT TAT       1056
Arg Asn Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr
            340                 345                 350

GCC ACC AGT GAT CTA ACA TTT GCT AAT GTT TCG TTT TCT TAT CCA AGC       1104
Ala Thr Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser
        355                 360                 365

AGA CCT TCG GAA GCA GTT TTA AAG AAC GTT AGT TTA AAT TTC TCT GCA       1152
Arg Pro Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala
    370                 375                 380

GGA CAA TTT ACT TTC ATA GTA GGA AAA TCA GGC TCA GGT AAA TCT ACA       1200
Gly Gln Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr
385                 390                 395                 400
```

FIG.4B

```
TTA TCC AAC TTA TTA TTA AGG TTC TAC GAT GGC TAT AAT GGA TCG ATA      1248
Leu Ser Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile
                405                 410                 415

TCT ATC AAT GGC CAC AAT ATC CAA ACA ATC GAC CAA AAA TTG CTA ATT      1296
Ser Ile Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile
            420                 425                 430

GAA AAT ATC ACC GTC GTA GAA CAG TTT TCC TGG ATT ATG CCT GGC ACC      1344
Glu Asn Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr
        435                 440 |               445
                        STE6 ---|--- CFTR

ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC      1392
Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr
    450                 455                 460

AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GAC ATC TCC AAG TTT      1440
Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe
465                 470                 475                 480

GCA GAG AAA GAC AAT ATA GTT CTT GGA GAA GGT GGA ATC ACA CTG AGT|     1488
Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser|
                485                 490                 495|
                                                    CFTR ---| STE6

GGC GGG CAA CAA CAA AGA GTT GCT ATA GCA CGT GCA TTC ATC AGA GAT      1536
Gly Gly Gln Gln Gln Arg Val Ala Ile Ala Arg Ala Phe Ile Arg Asp
            500                 505                 510

ACT CCA ATA TTA TTC TTA GAC GAA GCT GTA TCG GCT CTA GAT ATT GTT      1584
Thr Pro Ile Leu Phe Leu Asp Glu Ala Val Ser Ala Leu Asp Ile Val
        515                 520                 525

CAT CGC AAC CTG TTG ATG AAG GCA ATT AGG CAT TGG AGG AAA GGA AAG      1632
His Arg Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys
    530                 535                 540

ACT ACA ATC ATA TTG ACG CAT GAG TTG AGC CAA ATT GAA TCT GAT GAC      1680
Thr Thr Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp
545                 550                 555                 560

TAT TTA TAT TTA ATG AAG GAA GGT GAA GTT GTT GAA AGC GGC ACC CAG      1728
Tyr Leu Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln
                565                 570                 575

TCT GAA CTT CTA GCC GAT CCG ACC ACT ACA TTT AGC ACA TGG TAT CAC      1776
Ser Glu Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His
            580                 585                 590

CTA CAG AAT GAC TAC TCT GAT GCG AAA ACT ATT GTA GAT ACA GAG ACT      1824
Leu Gln Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr
        595                 600                 605
```

FIG. 4C

```
GAA GAA AAA TCT ATA CAC ACT GTG GAA AGT TTT AAC TCT CAA TTG GAA      1872
Glu Glu Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu
    610                 615                 620

ACA CCA AAA CTT GGA TCA TGC TTA AGT AAT CTG GGA TAT GAT GAG ACA      1920
Thr Pro Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr
625                 630                 635                 640

GAT CAG TTG TCC TTT TAC GAA GCA ATC TAT CAA AAA AGA TCG AAC GTT      1968
Asp Gln Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val
                645                 650                 655

AGA ACA AGA AGG GTT AAA GTT GAA GAG GAA AAT ATT GGG TAT GCA CTA      2016
Arg Thr Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu
            660                 665                 670

AAA CAA CAA AAG AAC ACC GAA AGT TCA ACA GGG CCA CAA CTT CTG AGC      2064
Lys Gln Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser
        675                 680                 685

ATT ATT CAG ATT ATC AAA AGA ATG ATT AAA AGC ATA AGA TAC AAA AAA      2112
Ile Ile Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys
    690                 695                 700

ATT CTA ATC TTG GGA CTG CTA TGT TCT CTT ATC GCA GGC GCC ACA AAT      2160
Ile Leu Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn
705                 710                 715                 720

CCC GTC TTT TCA TAC ACA TTC AGT TTC TTA CTA GAA GGA ATT GTC CCA      2208
Pro Val Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro
                725                 730                 735

TCC ACG GAT GGA AAA ACT GGC TCT TCA CAT TAT TTG GCG AAA TGG TCG      2256
Ser Thr Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser
            740                 745                 750

CTT CTT GTT CTT GGT GTG GCT GCG GCA GAT GGT ATT TTC AAT TTT GCT      2304
Leu Leu Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala
        755                 760                 765

AAA GGA TTC CTA TTA GAT TGC TGC AGT GAA TAC TGG GTT ATG GAT CTT      2352
Lys Gly Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu
    770                 775                 780

AGA AAT GAA GTT ATG GAA AAA CTG ACG AGA AAG AAT ATG GAC TGG TTT      2400
Arg Asn Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe
785                 790                 795                 800

TCT GGT GAA AAC AAC AAG GCT TCT GAA ATT TCT GCT CTA GTC TTG AAT      2448
Ser Gly Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn
                805                 810                 815
```

FIG. 4D

```
GAT TTG CGA GAT TTG AGG TCT TTG GTC TCT GAA TTT TTG AGT GCA ATG      2496
Asp Leu Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met
            820                 825                 830

ACT AGT TTC GTT ACC GTA TCA ACG ATT GGA CTA ATT TGG GCG TTA GTA      2544
Thr Ser Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val
        835                 840                 845

TCG GGC TGG AAG TTA AGT TTG GTT TGT ATT TCG ATG TTT CCA CTC ATA      2592
Ser Gly Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile
    850                 855                 860

ATT ATA TTT TCA GCA ATA TAT GGA GGC ATT TTA CAA AAA TGC GAA ACA      2640
Ile Ile Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr
865                 870                 875                 880

GAT TAT AAG ACA TCT GTT GCT CAG TTA GAA AAC TGC CTG TAC CAG ATT      2688
Asp Tyr Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile
            885                 890                 895

GTC ACT AAC ATT AAA ACC ATT AAG TGC TTA CAA GCT GAA TTT CAT TTT      2736
Val Thr Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe
            900                 905                 910

CAA TTG ACC TAC CAT GAC TTG AAG ATA AAA ATG CAA CAA ATT GCC TCC      2784
Gln Leu Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser
        915                 920                 925

AAA CGC GCC ATT GCC ACA GGA TTT GGT ATA TCT ATG ACA AAC ATG ATT      2832
Lys Arg Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile
    930                 935                 940

GTC ATG TGT ATC CAA GCT ATT ATT TAC TAC TAT GGC CTA AAG CTG GTT      2880
Val Met Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val
945                 950                 955                 960

ATG ATT CAC GAG TAC ACC TCA AAG GAA ATG TTT ACG ACT TTC ACT TTG      2928
Met Ile His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu
            965                 970                 975

TTA TTA TTC ACT ATT ATG TCA TGC ACT AGC CTA GTA AGT CAG ATA CCC      2976
Leu Leu Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro
            980                 985                 990

GAT ATA AGT AGA GGC CAA CGT GCT GCC AGT TGG ATC TAT AGG ATT CTT      3024
Asp Ile Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu
        995                 1000                1005

GAT GAA AAG CAT AAT ACC CTA GAG GTT GAA AAC AAT AAT GCT AGA ACA      3072
Asp Glu Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr
    1010                1015                1020
```

FIG. 4E

```
GTG GGA ATA GCT GGT CAC ACC TAC CAT GGC AAA GAA AAA AAA CCA ATC      3120
Val Gly Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile
1025                1030                1035                1040

GTT TCA ATT CAA AAT TTG ACA TTT GCC TAT CCA TCT GCA CCT ACC GCC      3168
Val Ser Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala
                    1045                1050                1055

TTT GTT TAC AAA AAC ATG AAT TTT GAC ATG TTT TGC GGA CAG ACG TTA      3216
Phe Val Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu
            1060                1065                1070

GGT ATC ATT GGT GAA TCA GGC ACA GGA AAG TCT ACA CTT GTG CTT TTA      3264
Gly Ile Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu
        1075                1080                1085

TTA ACA AAA CTT TAT AAT TGT GAA GTA GGC AAA ATT AAA ATA GAC GGT      3312
Leu Thr Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly
    1090                1095                1100

ACG GAC GTA AAT GAC TGG AAT TTG ACA AGT TTA AGA AAA GAA ATT TCA      3360
Thr Asp Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser
1105                1110                1115                1120

GTG GTT GAG CAA AAA CCT TTA TTA TTC AAT GGA ACC ATC AGA GAT AAC      3408
Val Val Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn
                    1125                1130                1135

CTA ACT TAT GGT TTA CAA GAT GAA ATA CTT GAA ATT GAA ATG TAT GAT      3456
Leu Thr Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp
            1140                1145                1150

GCA TTA AAA TAC GTA GGA ATC CAT GAC TTT GTA ATT TCA TCA CCT CAG      3504
Ala Leu Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln
        1155                1160                1165

GGC TTG GAT ACA CGT ATT GAT ACA ACT TTA CTG TCA GGT GGA CAA GCG      3552
Gly Leu Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala
    1170                1175                1180

CAA AGG CTT TGC ATA GCC AGA GCA CTT CTG AGA AAA TCA AAA ATT CTG      3600
Gln Arg Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu
1185                1190                1195                1200

ATT TTA GAT GAG TGT ACT TCA GCC TTG GAT TCT GTC AGC TCC TCT ATC      3648
Ile Leu Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile
                    1205                1210                1215

ATC AAT GAG ATC GTC AAA AAA GGT CCA CCT GCT CTA CTA ACA ATG GTT      3696
Ile Asn Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val
            1220                1225                1230
```

FIG. 4F

```
ATA ACG CAT AGT GAA CAA ATG ATG AGG TCT TGT AAC TCG ATT GCA GTT      3744
Ile Thr His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val
        1235                1240                1245

CTT AAA GAT GGT AAA GTG GTT GAG CGA GGT AAC TTC GAC ACT TTA TAT      3792
Leu Lys Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr
    1250                1255                1260

AAT AAT CGC GGG GAA TTA TTC CAA ATT GTT TCC AAC CAA AGC AGT TAA      3840
Asn Asn Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser *
1265                1270                1275                1280
```

FIG. 4G

H5 Chimera (Seq. ID No. 3)

```
ATG AAC TTT TTA AGT TTT AAG ACT ACA AAA CAC TAT CAC ATT TTC AGG          48
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
 1               5                  10                  15

TAC GTG AAC ATA CGG AAT GAC TAC AGG CTG TTA ATG ATA ATG ATA ATA          96
Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
            20                  25                  30

GGT ACC GTG GCA ACA GGC CTA GTG CCG GCA ATT ACT TCT ATC CTG ACG         144
Gly Thr Val Ala Thr Gly Leu Val Pro Ala Ile Thr Ser Ile Leu Thr
        35                  40                  45

GGC AGA GTG TTC GAT CTA CTA TCA GTT TTC GTG GCT AAT GGG TCA CAT         192
Gly Arg Val Phe Asp Leu Leu Ser Val Phe Val Ala Asn Gly Ser His
    50                  55                  60

CAA GGT TTG TAT TCC CAA CTA GTA CAG AGG TCA ATG GCA GTA ATG GCA         240
Gln Gly Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala
65                  70                  75                  80

CTT GGT GCG GCT TCT GTG CCA GTA ATG TGG CTT TCT CTA ACA AGT TGG         288
Leu Gly Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp
                85                  90                  95

ATG CAC ATC GGC GAG AGA CAA GGC TTT AGA ATA CGG TCA CAG ATA TTG         336
Met His Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu
            100                 105                 110

GAG GCA TAT TTG GAG GAA AAG CCA ATG GAA TGG TAC GAC AAT AAT GAA         384
Glu Ala Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu
        115                 120                 125

AAA TTG TTA GGA GAT TTT ACT CAA ATC AAC AGA TGT GTG GAA GAG CTA         432
Lys Leu Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu
    130                 135                 140

AGA TCA AGC TCC GCA GAG GCA TCA GCC ATA ACT TTC CAG AAT TTA GTT         480
Arg Ser Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val
145                 150                 155                 160

GCA ATA TGT GCG CTT CTG GGG ACG TCA TTC TAC TAT TCT TGG TCA TTA         528
Ala Ile Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu
                165                 170                 175

ACT TTA ATT ATT CTT TGC AGC TCT CCA ATA ATC ACA TTT TTT GCA GTG         576
Thr Leu Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val
            180                 185                 190
```

FIG.5A

```
GTG TTT TCC AGA ATG ATT CAT GTA TAT TCA GAG AAG GAG AAT TCT GAA        624
Val Phe Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu
        195                 200                 205

ACG AGT AAA GCA GCC CAA TTA CTT ACA TGG TCG ATG AAT GCC GCT CAA        672
Thr Ser Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln
    210                 215                 220

TTA GTG AGA TTA TAT TGT ACA CAA CGT CTA GAA AGG AAA AAA TTC AAG        720
Leu Val Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys
225                 230                 235                 240

GAA ATC ATA CTA AAT TGT AAC ACT TTC TTC ATC AAG AGT TGC TTT TTT        768
Glu Ile Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe
                245                 250                 255

GTT GCT GCA AAC GCT GGG ATC TTG AGA TTT TTG ACG TTG ACT ATG TTT        816
Val Ala Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe
            260                 265                 270

GTA CAG GGA TTC TGG TTT GGT TCC GCA ATG ATC AAA AAG GGC AAG CTG        864
Val Gln Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu
        275                 280                 285

AAC ATT AAC GAT GTA ATC ACT TGC TTC CAT TCA TGC ATT ATG CTG GGC        912
Asn Ile Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly
    290                 295                 300

TCG ACT TTA AAT AAT ACA TTA CAC CAA ATA GTT GTT CTT CAA AAA GGC        960
Ser Thr Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly
305                 310                 315                 320

GGA GTG GCT ATG GAA AAA ATC ATG ACT CTA TTA AAA GAT GGA TCC AAG       1008
Gly Val Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys
                325                 330                 335

CGA AAT CCT TTA AAT AAA ACT GTA GCC CAC CAA TTT CCA CTA GAT TAT       1056
Arg Asn Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr
            340                 345                 350

GCC ACC AGT GAT CTA ACA TTT GCT AAT GTT TCG TTT TCT TAT CCA AGC       1104
Ala Thr Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser
        355                 360                 365

AGA CCT TCG GAA GCA GTT TTA AAG AAC GTT AGT TTA AAT TTC TCT GCA       1152
Arg Pro Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala
    370                 375                 380

GGA CAA TTT ACT TTC ATA GTA GGA AAA TCA GGC TCA GGT AAA TCT ACA       1200
Gly Gln Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr
385                 390                 395                 400
```

FIG. 5B

```
TTA TCC AAC TTA TTA TTA AGG TTC TAC GAT GGC TAT AAT GGA TCG ATA      1248
Leu Ser Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile
                405                 410                 415

TCT ATC AAT GGC CAC AAT ATC CAA ACA ATC GAC CAA AAA TTG CTA ATT      1296
Ser Ile Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile
            420                 425                 430

GAA AAT ATC ACC GTC GTA GAA CAG TTT TCC TGG ATT ATG CCT GGC ACC      1344
Glu Asn Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr
        435                 440                 445
                    STE6 ---|--- CFTR

ATT AAA GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC      1392
Ile Lys Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr
    450                 455                 460

AGA AGC GTC ATC AAA GCA TGC CAA CTA GAA GAG GAC ATC TCC AAG TTT      1440
Arg Ser Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe
465                 470                 475                 480

GCA GAG AAA GAC AAT ATA GTT CTT GGA GAA GGT GGA ATC ACA CTG AGT      1488
Ala Glu Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser
                485                 490                 495

GGA GGT CAA CGA GCA AGA ATT TCT TTA GCA AGA GCA TTC ATC AGA GAT      1536
Gly Gly Gln Arg Ala Arg Ile Ser Leu Ala Arg Ala Phe Ile Arg Asp
            500                 505                 510
                            CFTR ---|--- STE6

ACT CCA ATA TTA TTC TTA GAC GAA GCT GTA TCG GCT CTA GAT ATT GTT      1584
Thr Pro Ile Leu Phe Leu Asp Glu Ala Val Ser Ala Leu Asp Ile Val
        515                 520                 525

CAT CGC AAC CTG TTG ATG AAG GCA ATT AGG CAT TGG AGG AAA GGA AAG      1632
His Arg Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys
    530                 535                 540

ACT ACA ATC ATA TTG ACG CAT GAG TTG AGC CAA ATT GAA TCT GAT GAC      1680
Thr Thr Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp
545                 550                 555                 560

TAT TTA TAT TTA ATG AAG GAA GGT GAA GTT GTT GAA AGC GGC ACC CAG      1728
Tyr Leu Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln
                565                 570                 575

TCT GAA CTT CTA GCC GAT CCG ACC ACT ACA TTT AGC ACA TGG TAT CAC      1776
Ser Glu Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His
            580                 585                 590

CTA CAG AAT GAC TAC TCT GAT GCG AAA ACT ATT GTA GAT ACA GAG ACT      1824
Leu Gln Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr
        595                 600                 605
```

FIG.5C

```
GAA GAA AAA TCT ATA CAC ACT GTG GAA AGT TTT AAC TCT CAA TTG GAA      1872
Glu Glu Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu
    610                 615                 620

ACA CCA AAA CTT GGA TCA TGC TTA AGT AAT CTG GGA TAT GAT GAG ACA      1920
Thr Pro Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr
625                 630                 635                 640

GAT CAG TTG TCC TTT TAC GAA GCA ATC TAT CAA AAA AGA TCG AAC GTT      1968
Asp Gln Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val
                645                 650                 655

AGA ACA AGA AGG GTT AAA GTT GAA GAG GAA AAT ATT GGG TAT GCA CTA      2016
Arg Thr Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu
            660                 665                 670

AAA CAA CAA AAG AAC ACC GAA AGT TCA ACA GGG CCA CAA CTT CTG AGC      2064
Lys Gln Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser
        675                 680                 685

ATT ATT CAG ATT ATC AAA AGA ATG ATT AAA AGC ATA AGA TAC AAA AAA      2112
Ile Ile Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys
    690                 695                 700

ATT CTA ATC TTG GGA CTG CTA TGT TCT CTT ATC GCA GGC GCC ACA AAT      2160
Ile Leu Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn
705                 710                 715                 720

CCC GTC TTT TCA TAC ACA TTC AGT TTC TTA CTA GAA GGA ATT GTC CCA      2208
Pro Val Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro
                725                 730                 735

TCC ACG GAT GGA AAA ACT GGC TCT TCA CAT TAT TTG GCG AAA TGG TCG      2256
Ser Thr Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser
            740                 745                 750

CTT CTT GTT CTT GGT GTG GCT GCG GCA GAT GGT ATT TTC AAT TTT GCT      2304
Leu Leu Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala
        755                 760                 765

AAA GGA TTC CTA TTA GAT TGC TGC AGT GAA TAC TGG GTT ATG GAT CTT      2352
Lys Gly Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu
    770                 775                 780

AGA AAT GAA GTT ATG GAA AAA CTG ACG AGA AAG AAT ATG GAC TGG TTT      2400
Arg Asn Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe
785                 790                 795                 800

TCT GGT GAA AAC AAC AAG GCT TCT GAA ATT TCT GCT CTA GTC TTG AAT      2448
Ser Gly Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn
                805                 810                 815
```

FIG.5D

```
GAT TTG CGA GAT TTG AGG TCT TTG GTC TCT GAA TTT TTG AGT GCA ATG      2496
Asp Leu Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met
            820                 825                 830

ACT AGT TTC GTT ACC GTA TCA ACG ATT GGA CTA ATT TGG GCG TTA GTA      2544
Thr Ser Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val
        835                 840                 845

TCG GGC TGG AAG TTA AGT TTG GTT TGT ATT TCG ATG TTT CCA CTC ATA      2592
Ser Gly Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile
    850                 855                 860

ATT ATA TTT TCA GCA ATA TAT GGA GGC ATT TTA CAA AAA TGC GAA ACA      2640
Ile Ile Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr
865                 870                 875                 880

GAT TAT AAG ACA TCT GTT GCT CAG TTA GAA AAC TGC CTG TAC CAG ATT      2688
Asp Tyr Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile
                885                 890                 895

GTC ACT AAC ATT AAA ACC ATT AAG TGC TTA CAA GCT GAA TTT CAT TTT      2736
Val Thr Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe
            900                 905                 910

CAA TTG ACC TAC CAT GAC TTG AAG ATA AAA ATG CAA CAA ATT GCC TCC      2784
Gln Leu Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser
        915                 920                 925

AAA CGC GCC ATT GCC ACA GGA TTT GGT ATA TCT ATG ACA AAC ATG ATT      2832
Lys Arg Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile
    930                 935                 940

GTC ATG TGT ATC CAA GCT ATT ATT TAC TAC TAT GGC CTA AAG CTG GTT      2880
Val Met Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val
945                 950                 955                 960

ATG ATT CAC GAG TAC ACC TCA AAG GAA ATG TTT ACG ACT TTC ACT TTG      2928
Met Ile His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu
                965                 970                 975

TTA TTA TTC ACT ATT ATG TCA TGC ACT AGC CTA GTA AGT CAG ATA CCC      2976
Leu Leu Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro
            980                 985                 990

GAT ATA AGT AGA GGC CAA CGT GCT GCC AGT TGG ATC TAT AGG ATT CTT      3024
Asp Ile Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu
        995                 1000                1005

GAT GAA AAG CAT AAT ACC CTA GAG GTT GAA AAC AAT AAT GCT AGA ACA      3072
Asp Glu Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr
    1010                1015                1020
```

FIG.5E

```
GTG GGA ATA GCT GGT CAC ACC TAC CAT GGC AAA GAA AAA AAA CCA ATC      3120
Val Gly Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile
1025                1030                1035                1040

GTT TCA ATT CAA AAT TTG ACA TTT GCC TAT CCA TCT GCA CCT ACC GCC      3168
Val Ser Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala
                1045                1050                1055

TTT GTT TAC AAA AAC ATG AAT TTT GAC ATG TTT TGC GGA CAG ACG TTA      3216
Phe Val Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu
            1060                1065                1070

GGT ATC ATT GGT GAA TCA GGC ACA GGA AAG TCT ACA CTT GTG CTT TTA      3264
Gly Ile Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu
        1075                1080                1085

TTA ACA AAA CTT TAT AAT TGT GAA GTA GGC AAA ATT AAA ATA GAC GGT      3312
Leu Thr Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly
    1090                1095                1100

ACG GAC GTA AAT GAC TGG AAT TTG ACA AGT TTA AGA AAA GAA ATT TCA      3360
Thr Asp Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser
1105                1110                1115                1120

GTG GTT GAG CAA AAA CCT TTA TTA TTC AAT GGA ACC ATC AGA GAT AAC      3408
Val Val Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn
                1125                1130                1135

CTA ACT TAT GGT TTA CAA GAT GAA ATA CTT GAA ATT GAA ATG TAT GAT      3456
Leu Thr Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp
            1140                1145                1150

GCA TTA AAA TAC GTA GGA ATC CAT GAC TTT GTA ATT TCA TCA CCT CAG      3504
Ala Leu Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln
        1155                1160                1165

GGC TTG GAT ACA CGT ATT GAT ACA ACT TTA CTG TCA GGT GGA CAA GCG      3552
Gly Leu Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala
    1170                1175                1180

CAA AGG CTT TGC ATA GCC AGA GCA CTT CTG AGA AAA TCA AAA ATT CTG      3600
Gln Arg Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu
1185                1190                1195                1200

ATT TTA GAT GAG TGT ACT TCA GCC TTG GAT TCT GTC AGC TCC TCT ATC      3648
Ile Leu Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile
                1205                1210                1215

ATC AAT GAG ATC GTC AAA AAA GGT CCA CCT GCT CTA CTA ACA ATG GTT      3696
Ile Asn Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val
            1220                1225                1230
```

FIG.5F

```
ATA ACG CAT AGT GAA CAA ATG ATG AGG TCT TGT AAC TCG ATT GCA GTT      3744
Ile Thr His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val
        1235                1240                1245

CTT AAA GAT GGT AAA GTG GTT GAG CGA GGT AAC TTC GAC ACT TTA TAT      3792
Leu Lys Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr
    1250                1255                1260

AAT AAT CGC GGG GAA TTA TTC CAA ATT GTT TCC AAC CAA AGC AGT TAA      3840
Asn Asn Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser *
1265                1270                1275                1280
```

FIG.5G

H6 Chimera (Seq. ID No. 5)

```
ATG AAC TTT TTA AGT TTT AAG ACT ACA AAA CAC TAT CAC ATT TTC AGG          48
Met Asn Phe Leu Ser Phe Lys Thr Thr Lys His Tyr His Ile Phe Arg
 1               5                  10                  15

TAC GTG AAC ATA CGG AAT GAC TAC AGG CTG TTA ATG ATA ATG ATA ATA          96
Tyr Val Asn Ile Arg Asn Asp Tyr Arg Leu Leu Met Ile Met Ile Ile
            20                  25                  30

GGT ACC GTG GCA CTA TCA GTT TTC GTG GCT AAT GGG TCA CAT CAA GGT         144
Gly Thr Val Ala Leu Ser Val Phe Val Ala Asn Gly Ser His Gln Gly
        35                  40                  45

TTG TAT TCC CAA CTA GTA CAG AGG TCA ATG GCA GTA ATG GCA CTT GGT         192
Leu Tyr Ser Gln Leu Val Gln Arg Ser Met Ala Val Met Ala Leu Gly
    50                  55                  60

GCG GCT TCT GTG CCA GTA ATG TGG CTT TCT CTA ACA AGT TGG ATG CAC         240
Ala Ala Ser Val Pro Val Met Trp Leu Ser Leu Thr Ser Trp Met His
65                  70                  75                  80

ATC GGC GAG AGA CAA GGC TTT AGA ATA CGG TCA CAG ATA TTG GAG GCA         288
Ile Gly Glu Arg Gln Gly Phe Arg Ile Arg Ser Gln Ile Leu Glu Ala
                85                  90                  95

TAT TTG GAG GAA AAG CCA ATG GAA TGG TAC GAC AAT AAT GAA AAA TTG         336
Tyr Leu Glu Glu Lys Pro Met Glu Trp Tyr Asp Asn Asn Glu Lys Leu
            100                 105                 110

TTA GGA GAT TTT ACT CAA ATC AAC AGA TGT GTG GAA GAG CTA AGA TCA         384
Leu Gly Asp Phe Thr Gln Ile Asn Arg Cys Val Glu Glu Leu Arg Ser
        115                 120                 125

AGC TCC GCA GAG GCA TCA GCC ATA ACT TTC CAG AAT TTA GTT GCA ATA         432
Ser Ser Ala Glu Ala Ser Ala Ile Thr Phe Gln Asn Leu Val Ala Ile
    130                 135                 140

TGT GCG CTT CTG GGG ACG TCA TTC TAC TAT TCT TGG TCA TTA ACT TTA         480
Cys Ala Leu Leu Gly Thr Ser Phe Tyr Tyr Ser Trp Ser Leu Thr Leu
145                 150                 155                 160

ATT ATT CTT TGC AGC TCT CCA ATA ATC ACA TTT TTT GCA GTG GTG TTT         528
Ile Ile Leu Cys Ser Ser Pro Ile Ile Thr Phe Phe Ala Val Val Phe
                165                 170                 175

TCC AGA ATG ATT CAT GTA TAT TCA GAG AAG GAG AAT TCT GAA ACG AGT         576
Ser Arg Met Ile His Val Tyr Ser Glu Lys Glu Asn Ser Glu Thr Ser
            180                 185                 190
```

FIG. 6A

```
AAA GCA GCC CAA TTA CTT ACA TGG TCG ATG AAT GCC GCT CAA TTA GTG        624
Lys Ala Ala Gln Leu Leu Thr Trp Ser Met Asn Ala Ala Gln Leu Val
        195                 200                 205

AGA TTA TAT TGT ACA CAA CGT CTA GAA AGG AAA AAA TTC AAG GAA ATC        672
Arg Leu Tyr Cys Thr Gln Arg Leu Glu Arg Lys Lys Phe Lys Glu Ile
    210                 215                 220

ATA CTA AAT TGT AAC ACT TTC TTC ATC AAG AGT TGC TTT TTT GTT GCT        720
Ile Leu Asn Cys Asn Thr Phe Phe Ile Lys Ser Cys Phe Phe Val Ala
225                 230                 235                 240

GCA AAC GCT GGG ATC TTG AGA TTT TTG ACG TTG ACT ATG TTT GTA CAG        768
Ala Asn Ala Gly Ile Leu Arg Phe Leu Thr Leu Thr Met Phe Val Gln
                245                 250                 255

GGA TTC TGG TTT GGT TCC GCA ATG ATC AAA AAG GGC AAG CTG AAC ATT        816
Gly Phe Trp Phe Gly Ser Ala Met Ile Lys Lys Gly Lys Leu Asn Ile
            260                 265                 270

AAC GAT GTA ATC ACT TGC TTC CAT TCA TGC ATT ATG CTG GGC TCG ACT        864
Asn Asp Val Ile Thr Cys Phe His Ser Cys Ile Met Leu Gly Ser Thr
        275                 280                 285

TTA AAT AAT ACA TTA CAC CAA ATA GTT GTT CTT CAA AAA GGC GGA GTG        912
Leu Asn Asn Thr Leu His Gln Ile Val Val Leu Gln Lys Gly Gly Val
    290                 295                 300

GCT ATG GAA AAA ATC ATG ACT CTA TTA AAA GAT GGA TCC AAG CGA AAT        960
Ala Met Glu Lys Ile Met Thr Leu Leu Lys Asp Gly Ser Lys Arg Asn
305                 310                 315                 320

CCT TTA AAT AAA ACT GTA GCC CAC CAA TTT CCA CTA GAT TAT GCC ACC       1008
Pro Leu Asn Lys Thr Val Ala His Gln Phe Pro Leu Asp Tyr Ala Thr
                325                 330                 335

AGT GAT CTA ACA TTT GCT AAT GTT TCG TTT TCT TAT CCA AGC AGA CCT       1056
Ser Asp Leu Thr Phe Ala Asn Val Ser Phe Ser Tyr Pro Ser Arg Pro
            340                 345                 350

TCG GAA GCA GTT TTA AAG AAC GTT AGT TTA AAT TTC TCT GCA GGA CAA       1104
Ser Glu Ala Val Leu Lys Asn Val Ser Leu Asn Phe Ser Ala Gly Gln
        355                 360                 365

TTT ACT TTC ATA GTA GGA AAA TCA GGC TCA GGT AAA TCT ACA TTA TCC       1152
Phe Thr Phe Ile Val Gly Lys Ser Gly Ser Gly Lys Ser Thr Leu Ser
    370                 375                 380

AAC TTA TTA TTA AGG TTC TAC GAT GGC TAT AAT GGA TCG ATA TCT ATC       1200
Asn Leu Leu Leu Arg Phe Tyr Asp Gly Tyr Asn Gly Ser Ile Ser Ile
385                 390                 395                 400
```

FIG.6B

```
AAT GGC CAC AAT ATC CAA ACA ATC GAC CAA AAA TTG CTA ATT GAA AAT      1248
Asn Gly His Asn Ile Gln Thr Ile Asp Gln Lys Leu Leu Ile Glu Asn
                405                 410                 415

ATC ACC GTC GTA GAA CAG TTT TCC TGG ATT ATG CCT GGC ACC ATT AAA      1296
Ile Thr Val Val Glu Gln Phe Ser Trp Ile Met Pro Gly Thr Ile Lys
            420                 425                 430
            STE6 ——— CFTR
GAA AAT ATC ATC TTT GGT GTT TCC TAT GAT GAA TAT AGA TAC AGA AGC      1344
Glu Asn Ile Ile Phe Gly Val Ser Tyr Asp Glu Tyr Arg Tyr Arg Ser
        435                 440                 445

GTC ATC AAA GCA TGC CAA CTA GAA GAG GAC ATC TCC AAG TTT GCA GAG      1392
Val Ile Lys Ala Cys Gln Leu Glu Glu Asp Ile Ser Lys Phe Ala Glu
    450                 455                 460

AAA GAC AAT ATA GTT CTT GGA GAA GGT GGA ATC ACA CTG AGT GGA GGT      1440
Lys Asp Asn Ile Val Leu Gly Glu Gly Gly Ile Thr Leu Ser Gly Gly
465                 470                 475                 480

CAA CGA GCA AGA ATT TCT TTA GCA AGA GCA GTA TAC AAA GAT GCT GAT      1488
Gln Arg Ala Arg Ile Ser Leu Ala Arg Ala Val Tyr Lys Asp Ala Asp
                485                 490                 495

TTG TAT TTA TTA GAC TCT CCT TTT GGA TAC CTA GAT ATT GTT CAT CGC      1536
Leu Tyr Leu Leu Asp Ser Pro Phe Gly Tyr Leu Asp Ile Val His Arg
            500                 505                 510
                                CFTR ——— STE6
AAC CTG TTG ATG AAG GCA ATT AGG CAT TGG AGG AAA GGA AAG ACT ACA      1584
Asn Leu Leu Met Lys Ala Ile Arg His Trp Arg Lys Gly Lys Thr Thr
        515                 520                 525

ATC ATA TTG ACG CAT GAG TTG AGC CAA ATT GAA TCT GAT GAC TAT TTA      1632
Ile Ile Leu Thr His Glu Leu Ser Gln Ile Glu Ser Asp Asp Tyr Leu
    530                 535                 540

TAT TTA ATG AAG GAA GGT GAA GTT GTT GAA AGC GGC ACC CAG TCT GAA      1680
Tyr Leu Met Lys Glu Gly Glu Val Val Glu Ser Gly Thr Gln Ser Glu
545                 550                 555                 560

CTT CTA GCC GAT CCG ACC ACT ACA TTT AGC ACA TGG TAT CAC CTA CAG      1728
Leu Leu Ala Asp Pro Thr Thr Thr Phe Ser Thr Trp Tyr His Leu Gln
                565                 570                 575

AAT GAC TAC TCT GAT GCG AAA ACT ATT GTA GAT ACA GAG ACT GAA GAA      1776
Asn Asp Tyr Ser Asp Ala Lys Thr Ile Val Asp Thr Glu Thr Glu Glu
            580                 585                 590

AAA TCT ATA CAC ACT GTG GAA AGT TTT AAC TCT CAA TTG GAA ACA CCA      1824
Lys Ser Ile His Thr Val Glu Ser Phe Asn Ser Gln Leu Glu Thr Pro
        595                 600                 605
```

FIG. 6C

```
AAA CTT GGA TCA TGC TTA AGT AAT CTG GGA TAT GAT GAG ACA GAT CAG      1872
Lys Leu Gly Ser Cys Leu Ser Asn Leu Gly Tyr Asp Glu Thr Asp Gln
    610                 615                 620

TTG TCC TTT TAC GAA GCA ATC TAT CAA AAA AGA TCG AAC GTT AGA ACA      1920
Leu Ser Phe Tyr Glu Ala Ile Tyr Gln Lys Arg Ser Asn Val Arg Thr
625                 630                 635                 640

AGA AGG GTT AAA GTT GAA GAG GAA AAT ATT GGG TAT GCA CTA AAA CAA      1968
Arg Arg Val Lys Val Glu Glu Glu Asn Ile Gly Tyr Ala Leu Lys Gln
                645                 650                 655

CAA AAG AAC ACC GAA AGT TCA ACA GGG CCA CAA CTT CTG AGC ATT ATT      2016
Gln Lys Asn Thr Glu Ser Ser Thr Gly Pro Gln Leu Leu Ser Ile Ile
            660                 665                 670

CAG ATT ATC AAA AGA ATG ATT AAA AGC ATA AGA TAC AAA AAA ATT CTA      2064
Gln Ile Ile Lys Arg Met Ile Lys Ser Ile Arg Tyr Lys Lys Ile Leu
        675                 680                 685

ATC TTG GGA CTG CTA TGT TCT CTT ATC GCA GGC GCC ACA AAT CCC GTC      2112
Ile Leu Gly Leu Leu Cys Ser Leu Ile Ala Gly Ala Thr Asn Pro Val
    690                 695                 700

TTT TCA TAC ACA TTC AGT TTC TTA CTA GAA GGA ATT GTC CCA TCC ACG      2160
Phe Ser Tyr Thr Phe Ser Phe Leu Leu Glu Gly Ile Val Pro Ser Thr
705                 710                 715                 720

GAT GGA AAA ACT GGC TCT TCA CAT TAT TTG GCG AAA TGG TCG CTT CTT      2208
Asp Gly Lys Thr Gly Ser Ser His Tyr Leu Ala Lys Trp Ser Leu Leu
                725                 730                 735

GTT CTT GGT GTG GCT GCG GCA GAT GGT ATT TTC AAT TTT GCT AAA GGA      2256
Val Leu Gly Val Ala Ala Ala Asp Gly Ile Phe Asn Phe Ala Lys Gly
            740                 745                 750

TTC CTA TTA GAT TGC TGC AGT GAA TAC TGG GTT ATG GAT CTT AGA AAT      2304
Phe Leu Leu Asp Cys Cys Ser Glu Tyr Trp Val Met Asp Leu Arg Asn
        755                 760                 765

GAA GTT ATG GAA AAA CTG ACG AGA AAG AAT ATG GAC TGG TTT TCT GGT      2352
Glu Val Met Glu Lys Leu Thr Arg Lys Asn Met Asp Trp Phe Ser Gly
    770                 775                 780

GAA AAC AAC AAG GCT TCT GAA ATT TCT GCT CTA GTC TTG AAT GAT TTG      2400
Glu Asn Asn Lys Ala Ser Glu Ile Ser Ala Leu Val Leu Asn Asp Leu
785                 790                 795                 800

CGA GAT TTG AGG TCT TTG GTC TCT GAA TTT TTG AGT GCA ATG ACT AGT      2448
Arg Asp Leu Arg Ser Leu Val Ser Glu Phe Leu Ser Ala Met Thr Ser
                805                 810                 815
```

FIG.6D

```
TTC GTT ACC GTA TCA ACG ATT GGA CTA ATT TGG GCG TTA GTA TCG GGC      2496
Phe Val Thr Val Ser Thr Ile Gly Leu Ile Trp Ala Leu Val Ser Gly
            820                 825                 830

TGG AAG TTA AGT TTG GTT TGT ATT TCG ATG TTT CCA CTC ATA ATT ATA      2544
Trp Lys Leu Ser Leu Val Cys Ile Ser Met Phe Pro Leu Ile Ile Ile
        835                 840                 845

TTT TCA GCA ATA TAT GGA GGC ATT TTA CAA AAA TGC GAA ACA GAT TAT      2592
Phe Ser Ala Ile Tyr Gly Gly Ile Leu Gln Lys Cys Glu Thr Asp Tyr
    850                 855                 860

AAG ACA TCT GTT GCT CAG TTA GAA AAC TGC CTG TAC CAG ATT GTC ACT      2640
Lys Thr Ser Val Ala Gln Leu Glu Asn Cys Leu Tyr Gln Ile Val Thr
865                 870                 875                 880

AAC ATT AAA ACC ATT AAG TGC TTA CAA GCT GAA TTT CAT TTT CAA TTG      2688
Asn Ile Lys Thr Ile Lys Cys Leu Gln Ala Glu Phe His Phe Gln Leu
                885                 890                 895

ACC TAC CAT GAC TTG AAG ATA AAA ATG CAA CAA ATT GCC TCC AAA CGC      2736
Thr Tyr His Asp Leu Lys Ile Lys Met Gln Gln Ile Ala Ser Lys Arg
            900                 905                 910

GCC ATT GCC ACA GGA TTT GGT ATA TCT ATG ACA AAC ATG ATT GTC ATG      2784
Ala Ile Ala Thr Gly Phe Gly Ile Ser Met Thr Asn Met Ile Val Met
        915                 920                 925

TGT ATC CAA GCT ATT ATT TAC TAC TAT GGC CTA AAG CTG GTT ATG ATT      2832
Cys Ile Gln Ala Ile Ile Tyr Tyr Tyr Gly Leu Lys Leu Val Met Ile
    930                 935                 940

CAC GAG TAC ACC TCA AAG GAA ATG TTT ACG ACT TTC ACT TTG TTA TTA      2880
His Glu Tyr Thr Ser Lys Glu Met Phe Thr Thr Phe Thr Leu Leu Leu
945                 950                 955                 960

TTC ACT ATT ATG TCA TGC ACT AGC CTA GTA AGT CAG ATA CCC GAT ATA      2928
Phe Thr Ile Met Ser Cys Thr Ser Leu Val Ser Gln Ile Pro Asp Ile
                965                 970                 975

AGT AGA GGC CAA CGT GCT GCC AGT TGG ATC TAT AGG ATT CTT GAT GAA      2976
Ser Arg Gly Gln Arg Ala Ala Ser Trp Ile Tyr Arg Ile Leu Asp Glu
            980                 985                 990

AAG CAT AAT ACC CTA GAG GTT GAA AAC AAT AAT GCT AGA ACA GTG GGA      3024
Lys His Asn Thr Leu Glu Val Glu Asn Asn Asn Ala Arg Thr Val Gly
        995                 1000                1005

ATA GCT GGT CAC ACC TAC CAT GGC AAA GAA AAA AAA CCA ATC GTT TCA      3072
Ile Ala Gly His Thr Tyr His Gly Lys Glu Lys Lys Pro Ile Val Ser
    1010                1015                1020
```

FIG. 6E

```
ATT CAA AAT TTG ACA TTT GCC TAT CCA TCT GCA CCT ACC GCC TTT GTT      3120
Ile Gln Asn Leu Thr Phe Ala Tyr Pro Ser Ala Pro Thr Ala Phe Val
1025                1030                1035                1040

TAC AAA AAC ATG AAT TTT GAC ATG TTT TGC GGA CAG ACG TTA GGT ATC      3168
Tyr Lys Asn Met Asn Phe Asp Met Phe Cys Gly Gln Thr Leu Gly Ile
                1045                1050                1055

ATT GGT GAA TCA GGC ACA GGA AAG TCT ACA CTT GTG CTT TTA TTA ACA      3216
Ile Gly Glu Ser Gly Thr Gly Lys Ser Thr Leu Val Leu Leu Leu Thr
            1060                1065                1070

AAA CTT TAT AAT TGT GAA GTA GGC AAA ATT AAA ATA GAC GGT ACG GAC      3264
Lys Leu Tyr Asn Cys Glu Val Gly Lys Ile Lys Ile Asp Gly Thr Asp
        1075                1080                1085

GTA AAT GAC TGG AAT TTG ACA AGT TTA AGA AAA GAA ATT TCA GTG GTT      3312
Val Asn Asp Trp Asn Leu Thr Ser Leu Arg Lys Glu Ile Ser Val Val
    1090                1095                1100

GAG CAA AAA CCT TTA TTA TTC AAT GGA ACC ATC AGA GAT AAC CTA ACT      3360
Glu Gln Lys Pro Leu Leu Phe Asn Gly Thr Ile Arg Asp Asn Leu Thr
1105                1110                1115                1120

TAT GGT TTA CAA GAT GAA ATA CTT GAA ATT GAA ATG TAT GAT GCA TTA      3408
Tyr Gly Leu Gln Asp Glu Ile Leu Glu Ile Glu Met Tyr Asp Ala Leu
                1125                1130                1135

AAA TAC GTA GGA ATC CAT GAC TTT GTA ATT TCA TCA CCT CAG GGC TTG      3456
Lys Tyr Val Gly Ile His Asp Phe Val Ile Ser Ser Pro Gln Gly Leu
            1140                1145                1150

GAT ACA CGT ATT GAT ACA ACT TTA CTG TCA GGT GGA CAA GCG CAA AGG      3504
Asp Thr Arg Ile Asp Thr Thr Leu Leu Ser Gly Gly Gln Ala Gln Arg
        1155                1160                1165

CTT TGC ATA GCC AGA GCA CTT CTG AGA AAA TCA AAA ATT CTG ATT TTA      3552
Leu Cys Ile Ala Arg Ala Leu Leu Arg Lys Ser Lys Ile Leu Ile Leu
    1170                1175                1180

GAT GAG TGT ACT TCA GCC TTG GAT TCT GTC AGC TCC TCT ATC ATC AAT      3600
Asp Glu Cys Thr Ser Ala Leu Asp Ser Val Ser Ser Ser Ile Ile Asn
1185                1190                1195                1200

GAG ATC GTC AAA AAA GGT CCA CCT GCT CTA CTA ACA ATG GTT ATA ACG      3648
Glu Ile Val Lys Lys Gly Pro Pro Ala Leu Leu Thr Met Val Ile Thr
                1205                1210                1215

CAT AGT GAA CAA ATG ATG AGG TCT TGT AAC TCG ATT GCA GTT CTT AAA      3696
His Ser Glu Gln Met Met Arg Ser Cys Asn Ser Ile Ala Val Leu Lys
            1220                1225                1230
```

FIG. 6F

```
GAT GGT AAA GTG GTT GAG CGA GGT AAC TTC GAC ACT TTA TAT AAT AAT      3744
Asp Gly Lys Val Val Glu Arg Gly Asn Phe Asp Thr Leu Tyr Asn Asn
        1235                1240                1245

CGC GGG GAA TTA TTC CAA ATT GTT TCC AAC CAA AGC AGT TAA              3786
Arg Gly Glu Leu Phe Gln Ile Val Ser Asn Gln Ser Ser *
    1250                1255                1260
```

FIG.6G

INTERNATIONAL SEARCH REPORT

Internat[redacted] Application No
PCT[redacted]94/04379

A. CLASSIFICATION OF SUBJECT MATTER
IPC 5 C12N15/62 C07K13/00 G01N33/50 C12N1/19

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 5 C07K G01N C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JOURNAL OF CELLULAR BIOCHEMISTRY Supplement 16F; April 3-16, 1992 page 15 see abstract no. U 116 ---	1,2,6-8, 13-15
A	MOLECULAR BIOLOGY OF THE CELL vol. 3 , September 1992 page 192A Y. MCCLENDON ET AL. 'Analysis of cystic fibrosis-related mutations in the yeast Ste6 protein' see abstract no. 1116 --- -/--	1,15

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 September 1994	26 -09- 1994

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Montero Lopez, B

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	CELL vol. 73, no. 2 , 23 April 1993 , CAMBRIDGE, NA US pages 335 - 346 JOHN L. TEEM ET AL. 'Identification of revertants for the cystic fibrosis deltaF508 mutation using STE6-CFTR chimeras in yeast' see summary see page 335, right column, paragraph 4 - page 336, right column, paragraph 2 see page 336, right column, paragraph 4 - page 338, left column, paragraph 1 see page 342, left column, paragraph 2 - right column, paragraph 1; figure 1	1-15
P,X	BIOCHEMISTRY AND CELL BIOLOGY vol. 71, no. 11-2 , December 1993 page AXXV JOHN L. TEEM ET AL. 'The construction of STE6/CFTR chimeric transporters to define functional similarities between transmembrane domain segments' see abstract	1,7, 13-15

1